APES-USA

Academic
Performance
Evaluation of
Students
-
Ubiquitous
System
Analyzed

Letter Grading System is inherently Unfair by its very Design and requires a complete re-design

The Problem is not Grade Inflation

Prof. Keshava PRASAD Halemane, Ph.D.

ISBN:xxxxxxxxxx
ISBN-13:xxx-xxxxxxxxxx

DEDICATION

ಅಜ್ಜಿ (grandma) & ಅಜ್ಜ (grandpa)

ಅಮ್ಮ (mom) & ಅಪ್ಪ (dad)

for their teachings through love, that

in life '*quality*' matters more than '*quantity*' in the long run;

my wife & kids for their love;

and

'SQUIDS'

for the enlightening encounter on the beach on that evening
on Thursday the Eighth of August Two Thousand Two.

--- ಕೇಶವ ಪ್ರಸಾದ ಹಳೆಮನೆ.

ACKNOWLEDGMENTS

The author wishes to express his grateful appreciation and indebtedness to his Alma Mater as K.R.E.C. Surathkal, which is also his present employer as N.I.T.K. Surathkal, for providing the essential supportive environment.

CONTENTS

LIST OF TABLES & FIGURES

PREFACE

The Letter-Grading System (LGS) is the most imposing, impressive and widely prevalent approach adopted by the academia, for the purpose of Academic Performance Evaluation of Students, in most countries around the world. Here is a detailed report of a systematic study of the LGS system. It is shown conclusively that the LGS has some serious lacunae inherent in its very design, and because of which it fails to meet the essential purpose for which the *Academic Performance Evaluation of Students* (APES) is undertaken in Academia.

In the first part of the book, it is shown that the LGS system exhibits a highly chaotic behavior. Because of the system-intrinsic chaotic phenomena the LGS system is grossly *unfair to the students*, who get subjected to *chaotic system biased unfair and unreliable comparisons*. It is *unfair to the prospective employers* who expect some relevant, unbiased, reliable information to be conveyed through those grade reports. It is *unfair to the teachers*, who find themselves utterly helpless, when the raw-scores/marks that they had originally awarded, are later subjected to chaotic system biased information loss/corruption. Also, the LGS system design wrongly presumes that the teacher's precision in evaluation is rather poor, limited to classification into possibly only a handful of distinct categories. Interestingly, there is a *mathematical fallacy* that is intrinsically implied in the LGS system design philosophy itself, since a significantly higher precision level is mysteriously presumed to have been achieved in reporting the final numbers of GPA or CGPA, etc. The LGS system just simply fails to provide a reliable mechanism for a true representation and communication of the otherwise appropriate/relevant information as to what the teachers really meant to convey, regarding the measurements originally conducted, towards the academic performance evaluation of students, that is (and should be)

unquestionably considered as an integral component of the overall *Teaching-Learning-Evaluation-Review* Model Academic *Environment*.

The second part presents a detailed and systematic analysis of the chaotic system behavior. The system-intrinsic phenomena of chaotic system biased information loss/corruption is shown to be caused by the quantization and the contraction-expansion mapping that are implied in the very design of the LGS system and the associated computational model used for conversion of raw-scores to grade-points/letter-grades, GPA, CGPA, etc. Specifically, three kinds of chaotic system-intrinsic phenomena have been identified: Chaotic System Biased Amplification or Attenuation of Differentials, Chaotic System Biased Suppression or Expression of Differentials, and Chaotic System Biased Relative Rank Inversion. The root cause for all the undesirable system characteristics is the use of a very crude grade points scale with an unacceptably poor precision level, considering the anticipated/projected requirement of a far higher precision level in the GPA and/or CGPA figures to be reported in the grade reports. A possible mathematical system model is briefly visited, although a more rigorous presentation is beyond the scope of this book.

The third part presents the details of the "Student's Academic Performance Evaluation System" (SAPES) as a proposed solution to overcome this 'APES-USA' problem.

You may possibly feel uncomfortable with some of the terms and phrases, particularly when first encountered. Just keep reading, without feeling intimidated in any way, and you will soon discover that all the ideas and concepts surely fit together very well into a holistic systems approach to the study of the problem concerned.

--- Keshava PRASAD Halemane.
N.I.T.K.Surathkal, India.

Notable quotes - Quotable Notes

An academic curriculum is
for an educational system,
as a governing constitution is
for a national governance.

An evaluation system is to an academic curriculum,
as a judicial system is to a governing constitution.

Evaluation is a very critical step/phase
in any human intellectual endeavor.

Every evaluation is by and for comparison,
whether one likes it or not.

Every evaluation is a relative evaluation - either
relative in comparison with a standard, or maybe
relative in comparison with those being evaluated.

An evaluation system must meet
the fundamental quality characteristics of being
fair/justifiable, objective/unbiased, reliable/precise,
as well as robust/resilient.

Any computed (aggregate) quantity
cannot have a better precision than that of
the input data-set used there-for.

PART - I

The Problem

LETTER GRADING SYSTEM IS INHERENTLY UNFAIR

1.1 Introduction

The most imposing, impressive and widely prevalent *Letter Grading System* (LGS) model requires each of the teachers as evaluators, to assign an appropriately chosen letter grade (from among the set of assignable letter grades as specified in the academic rules & regulations and as per the guidelines laid out by the school (college or university or institute) to each of the students in each of the courses (subjects or papers) and report the same, at the end of every academic session (term or quarter or semester). For this purpose, the teacher may be required to conduct a sequence of tests (and/or quizzes and/or examinations, etc.), an almost continuous, incremental as well as cumulative evaluations, that together facilitate in assessing (evaluating or measuring) the extent (or degree) of proficiency achieved by the student in the subject, based on the performance of the student in such tests etc., which when well designed would evenly cover the entire subject material that a student is expected to master during that period of study.

This first part of the book contains a detailed study of the existing LGS system behavior for the Academic Performance Evaluation of Students. It is shown that the LGS system exhibits a chaotic (unpredictable although deterministic), highly complex, counter-intuitive as well as undesirable behavior. Specifically, three kinds of chaotic system-intrinsic phenomena have been identified:

Chaotic System Biased Amplification or Attenuation of Differentials (CSBAAD), Chaotic System Biased Suppression or Expression of Differentials (CSBSED), Chaotic System Biased Relative Rank Inversion (CSBRRI). These system-intrinsic phenomena of Chaotic System Biased Information Loss/Corruption (CSBILC) result in Chaotic System Biased Unfair and Unreliable Comparisons (CSBUUC) among the students being evaluated, making it to be very grossly unfair to the students. It is also unfair to the prospective employers who expect some relevant, unbiased, reliable information to be contained in and conveyed through those transcripts or grade reports. Again, it is very unfair to the teachers, who find themselves utterly helpless, when the Raw Scores/Marks that they had originally awarded very conscientiously, are later subjected to chaotic system biased information loss/corruption.

Also, the LGS system design wrongly presumes that the teacher's precision in assessment/evaluation is rather poor, limited to classification into possibly only a handful of distinct categories. However, there is a *mathematical fallacy* that is intrinsically implied in the LGS system design philosophy itself, since a significantly higher precision level is mysteriously presumed to have been achieved in reporting the final numbers of GPA or CGPA, etc. This is like expecting a double precision accuracy in some computational result when all the computations are performed in single precision although the input is in double precision. On the other hand, please note that *if in fact the GPA and CGPA can be reported after duly rounding them off to the grade point values corresponding to the nearest letter grade, then*

there is nothing much to claim by this study! However, we know very well that it cannot be so!

The LGS system just simply fails to provide a reliable/precise and robust/resilient mechanism for a true representation and communication of otherwise appropriate/relevant information as to what the teachers really meant to convey, regarding their unbiased/objective and fair/justifiable assessments/evaluations originally conducted.

1.2 A Typical Letter Grading System (LGS) Model

Although there are several minor variations among the different letter grading systems that are prevalent in the various educational institutions, in terms of the parameters incorporated in such a system design, the essential system behavior observed in these different scenarios happen to be the same! The observed differences and variations among them are only in terms of the nature and/or the extent and/or the locus, of the very same, essentially general, system behavior, which we present here, by taking a typical grading system model, although the same observations/comments can be made with regard to any other specific letter grading system model, belonging to the wide spectrum of various possible distinctly different specific system models.

Letter Grade	Grade Point	Lower Cut-off Mark	Upper Cut-off Mark
AA	10	[90	100]
AB	9	[80	90)
BB	8	[70	80)
BC	7	[60	70)
CC	6	[50	60)
CD	5	[40	50)
DD	4	[30	40)
FF	0	[00	30)

Table-1.01: A typical 10-point grade points scale,
with lower and upper cut-off marks

Letter Grade	Grade Point	Lower Cut-off Mark	Upper Cut-off Mark
A	5	[90	100]
B	4	[80	90)
C	3	[70	80)
D	2	[60	70)
E	1	[50	60)
F	0	[00	50)

Table-1.02: A typical 5-point grade points scale,
with lower and upper cut-off marks

Let us suppose that the set of valid assignable *Letter Grades* and the associated *Quality/Grade Points* are given. Table-1.01 gives a typical example with a 10-point grade points scale and Table-1.02 gives a typical example with a 5-point grade points scale.

That is, with a 10-point scale, the letter grades and the associated grade points may for example be as -
AA=10, AB=9, BB=8, BC=7, CC=6, CD=5, DD=4, FF(fail)=0;
whereas with a 5-point scale, the letter grades and the associated grade points may for example be as -
A=5, B=4, C=3, D=2, E=1, F(fail)=0.

Other letter grades (like Pass/Fail, etc) may not carry Quality/Grade Points, but only indicate certain essential information about the Academic Progress of a student towards one's Graduation Requirements. Once a letter grade is assigned to each of the courses that a student has "*registered for credit*", the *Grade Point Average* (GPA) for that student is computed as the weighted average of the *Grade Points* (corresponding to the letter grades) with the *Course Credits/Units* (CCU) of the corresponding courses as the multiplicative weighting factors for this computation.

Now let us focus on the problem that a typical teacher would have to face, in order to assign an appropriate letter grade to a student on a particular subject. The teacher does go through the evaluation of the answer scripts/papers of a sequence of tests etc, in order to determine the final letter grade, based on some well defined mechanism as per institute academic regulations, usually announced

to a class of students at the very beginning of the academic term. For example, a teacher may decide and therefore announce to the students that there would be four one-hour tests each carrying 25% weightage in determining the final grade. Now, how is a test paper evaluated? How, in a given test paper, the answer to each of the questions, evaluated? And how are these combined together to form what we call as the Raw-Score for a test paper?

We believe that no teacher starts off with awarding appropriately chosen letter grades to each of the various questions in a test. Usually, while setting the Question Paper for a test, appropriate numerical marks/scores are assigned for each of the questions, maybe on a Numerical Scale in the range 0-100 (Percentage Scale). While evaluating a test answer paper, appropriate mark/score is given for the answer to each of the questions, based on how close it is to the expected (standard?) answer, or to what extent the answer is acceptable (in case of possible multiple correct/acceptable answers, like in design problems), and these numerical scores are simply added together to come up with a total numerical score, which is the Raw-Score for that particular test paper.

If in fact it be possible, to assign letter grades right at the level of evaluation of each of the questions in a test, and if in fact there is a well accepted standard procedure designed to somehow combine them into an aggregate letter grade for that particular test, and then again combine (aggregation process) such letter grades for each of the four or more

tests/etc into a final letter grade for that student in that specific course; then we do not have much to say in this paper! The whole problem arises because one needs to resort to numbers, in order to have a well defined rational scheme for any measurements, and also to enable us in combining these various scores into a consolidated overall aggregate measure of whatever attribute is being measured.

It is quite evident that a teacher has to use some mapping-scheme to convert the raw-scores to the letter-grades (on a grade points scale). This mapping scheme may be either specified by the school or each teacher may be given the flexibility/freedom of designing one's own (although such ad hoc freedom, in itself, without any guidelines thereof, can possibly become an even more serious lacuna in the system). Any specification of a mapping scheme must include an appropriate partitioning of the raw scores scale into a sequence of intervals, so that each interval gets mapped onto its corresponding point in the grade point scale and therefore the corresponding letter grade. Although there can be several specific mapping schemes, they may be anywhere in between the two extreme generic types: (i) *distribution based mapping* and (ii) *distribution independent mapping*; thus giving rise to the corresponding classification of the grading system.

Distribution based grading system can be further classified into two broad categories - one of them may be called '*grading on the curve*' or '*grading based on some chosen standard distribution*' wherein the letter grades are *force-fitted* onto a standard normalized

distribution curve, like the Gaussian Probability Distribution Curve. Since the actual distribution of raw scores in a class is (supposed to be somewhat) independent of the teacher and the evaluation system, the only way that the letter grades can be made to fit into a given standard distribution is to have the mapping scheme dependent on the actual distribution of the raw scores in the class; hence this particular name for it. This approach has several proponents and/or followers. However, we believe that *any force-fitting performed on any set of raw data leads to Information Loss/Corruption,* and "we can only end up seeing what we look for". Therefore, for the purpose of having a typical grading system model, we shall not use this scheme. This decision in itself, is not a limitation (of being seemingly dependent on the system model) in terms of the scope of validity or the conclusions that can be drawn from this study, since any mapping scheme belonging to any of these two extreme types, or any other type, generic or otherwise, can be shown to have the *same common system characteristics* which in fact cause the system to exhibit the undesirable behavior that is explained in this paper.

A second sub-category under the generic type of distribution based grading system adapts its mapping scheme dependent upon the actual distribution of the raw scores in a class, and justifies that label, although there is no force-fitting of the letter grades (grade points) done to conform to any specific standard normalized distribution curve as such.

Distribution independent grading system uses a fixed mapping scheme that is based on some well defined divisions of the raw scores scale into intervals, the specification of these intervals, and also the associated mapping scheme being independent of the actual distribution of the raw scores in a class.

For the purpose of our study, a distribution independent mapping scheme is considered, as given in Table-1.01 for the case of a 10-point grade points scale, and as given in Table-1.02 for the case of a 5-point grade points scale.

That is :
[100-90]:AA,(90-80]:AB,(80-70]:BB,(70-60]:BC,(60-50]:CC,(50-40]:CD,(40-30]:DD,(30-00]:FF(fail)
with the 10-point scale, and

[100-90]: A,(90-80]: B,(80-70]: C,(70-60]: D,(60-50]: E, (50-00]:F (fail)
with the 5-point scale.

This mapping scheme, to convert the *Numerical Raw Scores* to *Letter Grades*, along with the above defined *Grade Points* for each of the letter grades, can be seen to result in an overall computational procedure for our typical system model; in order to determine the letter grade for each of the courses, for each of the students, using the available numerical raw scores; and also to compute the *Grade Point Average* (GPA), at the end of an academic term/semester; and similarly the Cumulative Grade Point Average (CGPA) for successive academic terms in the program of study. The parameters that define or *characterize* such a mapping scheme from the raw scores to the letter grades, and thus determine the associated computational procedure, include the raw-score interval-domains

(that gets mapped onto each of the letter-grades/Grade-Points) that are non-overlapping and laid out in a well ordered sequence. This characterization is common to any mapping scheme irrespective of whether it belongs to either of the two extremes, namely the "distribution based grading system" or "distribution independent grading system", or even any other possible LGS system design. The differences are in terms of the size and hence the relative positioning of each of the raw-score interval-domains that get mapped onto each of the letter-grades (grade point values); all other characteristics of the system behavior being essentially same. This justifies the above choice of a simple fixed mapping scheme for the sake of illustrating the system characteristics.

1.3 Chaotic LGS System Behavior

First, let us make a simple observation regarding the situation encountered in comparing the grades obtained by students in a specific single course. It is not at all uncommon that the raw score of one student will be at the lower end of the interval corresponding to one's grade while that of another will be at the high end of the corresponding interval. For example, consider a situation with four students S1, S2, S3, S4, with their raw scores being 99+, 90, 89+ and 80 respectively, thus getting their grades to be AA, AA, AB and AB respectively. A small difference of 1% (90 for S2, 89 for S3) in the raw scores can result in a difference of 10% (10 for S2, 9 for S3) in grade points, because their raw scores happen to be across an interval

boundary defined by the cut off marks of 90 between AA and AB. However, a large difference of almost 10% (between S1 & S2, or between S3 & S4) does not result in any difference in their grade points / grades. Minor differences in raw scores get unduly amplified in the corresponding grades (grade points) just because they are located across an interval boundary, while major differences in raw scores get unduly suppressed in the corresponding grades (grade points) just because they are located within the very same interval. This undesirable system behavior gets further chaotic when the grades of various courses are combined to compute the GPA and/or CGPA, etc, as will be evident from the following discussion.

Now, with the 10-point grade points scale, consider a scenario with a group of 32 students, with a wide range of academic performance levels; identified as 32S01, 32S02, . . . 32S32, each taking the six courses, of which three (s1/4c, s2/4c, s3/4c) are of 4 CCU, two (s4/3c, s5/3c) are of 3 CCU, and one (s6/2c) is of 2 CCU. Suppose that at the end of their academic term, their Raw Scores and Letter Grades in each of the courses are as listed in Table-1.03; and that Figure-1.1 gives the associated plot of sGPCCU versus sRSCCU for this case.

With the 5-point grade points scale, consider a scenario with a group of 22 students, again with a wide range of academic performance levels, identified as 22S01, 22S02, . . . 22S22, each taking the six courses, of which three (s1/4c, s2/4c, s3/4c) are of 4 CCU, two (s4/3c, s5/3c) are of 3 CCU, and one (s6/2c) is of 2 CCU.

Suppose that at the end of their academic term, their Raw Scores and Letter Grades in each of the courses are as listed in Table-1.04; and that Figure-1.2 gives the associated plot of sGPCCU versus sRSCCU for this case.

The Students Academic Performance Index (SAPI) is computed as follows: First, the sum (sRSCCU) of the raw scores, each of them being multipled by the corresponding Course Credits/Units, is computed. Then, sRSCCU is divided by the sum (sCCU) of Course Credits/Units (20, here) registered by the student for the semester; the ratio is the required SAPI value. To compute the Grade Point Average (GPA) first the sum (sGPCCU) of the Grade Points (associated with the Letter Grades) each of them being multiplied by the corresponding Course Credits/Units, is computed. Then, sGPCCU is divided by the sum (sCCU) of Course Credits/Units (20, here) registered by the student for the semester; the ratio is the required GPA value.

Table-1.03 along with Figure-1.1 gives, for each of the 32 students [Table-1.04 along with Figure-1.2 in case of the 5-point scale, for a group of 22 students], for each of the six courses, the Raw Score (RS) and the Letter Grade (LG), along with sRSCCU, SAPI, SAPI-Rank, and also sGPCCU, GPA, GPA-Rank. The computations are as explained above.

The above observations do reveal some bare naked truth, about the lacunae or weaknesses or drawbacks in the prevailing letter

grading system. Specifically, the LGS system exhibits three types of system-intrinsic chaotic behavior as listed in Table-1.05 [refer Table-1.06 for the 5-point scale].

One can observe from Figure-1.1 [Figure-1.2 in the case of the 5-point scale] that the computation of GPA (and/or CGPA) yields results that lie within a finite interval; that is, for each possible value of sGPCCU (or GPA or CGPA) there exists a finite interval of corresponding values of sRSCCU (or SAPI) and similarly for each possible value of sRSCCU (or SAPI) there exists a finite interval of corresponding values of sGPCCU (or GPA or CGPA).

Student ID	Course No. / Credit Units Raw-Score | Letter-Grade												sGPCCU	GPA	GPA-Rank	sRSCCU	SAPI	SAPI-Rank
	s1 / 4c		s2 / 4c		s3 / 4c		s4 / 3c		s5 / 3c		s6 / 2c							
	RS	LG	RS	LG	RS	LG	RS	LG	RS	LG	RS	LG						
32S01	99	A	99	A	99	A	99	A	99	A	99	A	200	10	1	1980	99	1
32S02	90	A	90	A	90	A	90	A	90	A	90	A	200	10	1	1800	90	3
32S03	90	A	90	A	90	A	90	A	90	A	80	A	198	9.9	3	1780	89	4
32S04	99	A	89	A	89	A	89	A	89	A	89	A	184	9.2	4	1820	91	2
32S05	80	A	80	A	80	A	80	A	80	A	90	A	182	9.1	5	1620	81	6
32S06	89	A	89	A	89	A	89	A	89	A	89	A	180	9	6	1780	89	4
32S07	80	A	80	A	80	A	80	A	80	A	80	A	180	9	6	1600	80	8
32S08	80	A	80	A	80	A	80	A	80	A	70	B	178	8.9	8	1580	79	9
32S09	89	A	79	B	79	B	79	B	79	B	79	B	164	8.2	9	1620	81	6
32S10	70	B	70	B	70	B	70	B	70	B	80	A	162	8.1	10	1420	71	11
32S11	79	B	79	B	79	B	79	B	79	B	79	B	160	8	11	1580	79	9
32S12	70	B	70	B	70	B	70	B	70	B	70	B	160	8	11	1400	70	13
32S13	70	B	70	B	70	B	70	B	70	B	60	B	158	7.9	13	1380	69	14
32S14	79	B	69	B	69	B	69	B	69	B	69	B	144	7.2	14	1420	71	11
32S15	60	B	60	B	60	B	60	B	60	B	70	B	142	7.1	15	1220	61	16
32S16	69	B	69	B	69	B	69	B	69	B	69	B	140	7	16	1380	69	14
32S17	60	B	60	B	60	B	60	B	60	B	60	B	140	7	16	1200	60	18
32S18	60	B	60	B	60	B	60	B	60	B	50	C	138	6.9	18	1180	59	19
32S19	69	B	59	C	59	C	59	C	59	C	59	C	124	6.2	19	1220	61	16
32S20	50	C	50	C	50	C	50	C	50	C	60	B	122	6.1	20	1020	51	21
32S21	59	C	59	C	59	C	59	C	59	C	59	C	120	6	21	1180	59	19
32S22	50	C	50	C	50	C	50	C	50	C	50	C	120	6	21	1000	50	23
32S23	50	C	50	C	50	C	50	C	50	C	40	C	118	5.9	23	980	49	24
32S24	59	C	49	C	49	C	49	C	49	C	49	C	104	5.2	24	1020	51	21
32S25	40	C	40	C	40	C	40	C	40	C	50	C	102	5.1	25	820	41	26
32S26	49	C	49	C	49	C	49	C	49	C	49	C	100	5	26	980	49	24
32S27	40	C	40	C	40	C	40	C	40	C	40	C	100	5	26	800	40	28
32S28	40	C	40	C	40	C	40	C	40	C	30	D	98	4.9	28	780	39	29
32S29	49	C	39	D	39	D	39	D	39	D	39	D	84	4.2	29	820	41	26
32S30	30	D	30	D	30	D	30	D	30	D	40	C	82	4.1	30	620	31	31
32S31	39	D	39	D	39	D	39	D	39	D	39	D	80	4	31	780	39	29
32S32	30	D	30	D	30	D	30	D	30	D	30	D	80	4	31	600	30	32

Table 1.03: A group of 32 students with wide-ranging performance levels (10-point grading scale)

Student ID	Course No. / Credit Units Raw-Score | Letter-Grade												sGPCCU	GPA	GPA-Rank	sRSCCU	SAPI	SAPI-Rank
	s1 / 4c		s2 / 4c		s3 / 4c		s4 / 3c		s5 / 3c		s6 / 2c							
	RS	LG	RS	LG	RS	LG	RS	LG	RS	LG	RS	LG						
22S01	99	A	99	A	99	A	99	A	99	A	99	A	100	5	1	1980	99	1
22S02	90	A	90	A	90	A	90	A	90	A	90	A	100	5	1	1800	90	3
22S03	90	A	90	A	90	A	90	A	90	A	80	B	98	4.9	3	1780	89	4
22S04	99	A	89	B	89	B	89	B	89	B	89	B	84	4.2	4	1820	91	2
22S05	80	B	80	B	80	B	80	B	80	B	90	A	82	4.1	5	1620	81	6
22S06	89	B	89	B	89	B	89	B	89	B	89	B	80	4	6	1780	89	4
22S07	80	B	80	B	80	B	80	B	80	B	80	B	80	4	6	1600	80	8
22S08	80	B	80	B	80	B	80	B	80	B	70	C	78	3.9	8	1580	79	9
22S09	89	B	79	C	79	C	79	C	79	C	79	C	64	3.2	9	1620	81	6
22S10	70	C	70	C	70	C	70	C	70	C	80	B	62	3.1	10	1420	71	11
22S11	79	C	79	C	79	C	79	C	79	C	79	C	60	3	11	1580	79	9
22S12	70	C	70	C	70	C	70	C	70	C	70	C	60	3	11	1400	70	13
22S13	70	C	70	C	70	C	70	C	70	C	60	D	58	2.9	13	1380	69	14
22S14	79	C	69	D	69	D	69	D	69	D	69	D	44	2.2	14	1420	71	11
22S15	60	D	60	D	60	D	60	D	60	D	70	C	42	2.1	15	1220	61	16
22S16	69	D	69	D	69	D	69	D	69	D	69	D	40	2	16	1380	69	14
22S17	60	D	60	D	60	D	60	D	60	D	60	D	40	2	16	1200	60	18
22S18	60	D	60	D	60	D	60	D	60	D	50	E	38	1.9	18	1180	59	19
22S19	69	D	59	E	59	E	59	E	59	E	59	E	24	1.2	19	1220	61	16
22S20	50	E	50	E	50	E	50	E	50	E	60	D	22	1.1	20	1020	51	21
22S21	59	E	59	E	59	E	59	E	59	E	59	E	20	1	21	1180	59	19
22S22	50	E	50	E	50	E	50	E	50	E	50	E	20	1	21	1000	50	22

Table 1.04: A group of 22 students with wide-ranging performance levels (5-point grading scale)

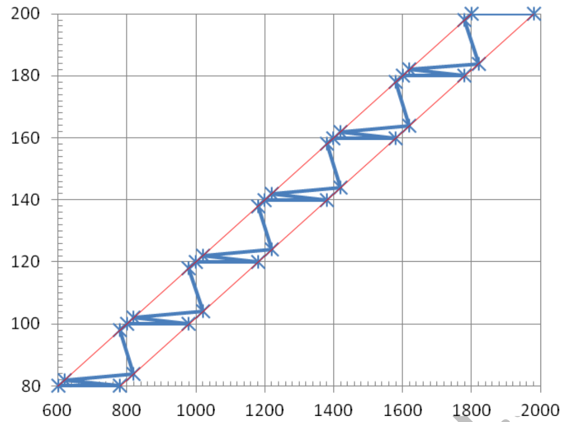

Figure-1.1: Plot of sGPCCU versus sRSCCU for 32 students
(10-point grading scale)

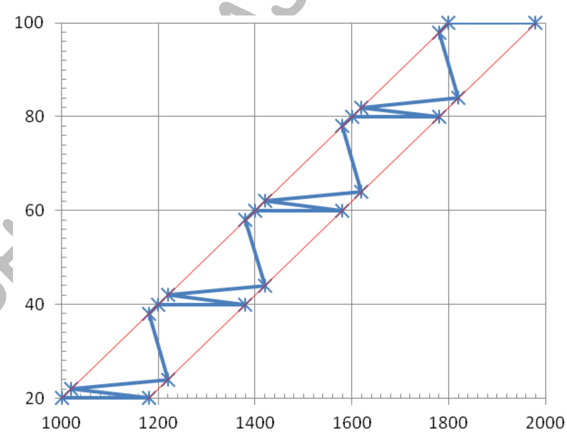

Figure-1.2: Plot of sGPCCU versus sRSCCU for 22 students
(5-point grading scale)

(1) Phenomenon of Chaotic System Biased Amplification or Attenuation of Differentials (CSBAAD):

(a) Small difference in GPA associated with large difference in SAPI:
(32S01-32S03), (32S04-32S05), (32S06-32S08), (32S09-32S10),
(32S11-32S13), (32S14-32S15), (32S16-32S18), (32S19-32S20),
(32S21-32S23), (32S24-32S25), (32S26-32S28), (32S29-32S30);

(b) Large difference in GPA associated with small difference in SAPI:
(32S02-32S06), (32S07-32S11), (32S12-32S16),
(32S17-32S21), (32S22-32S26), (32S27-32S31);

---

(2) Phenomenon of Chaotic System Biased Suppression or Expression of Differentials (CSBSED):

(a) Large difference in SAPI associated with no difference in GPA:
(32S01-32S02), (32S06-32S07), (32S11-32S12), (32S16-32S17),
(32S21-32S22), (32S26-32S27), (32S31-32S32);

(b) No difference in SAPI associated with large difference in GPA:
(32S03-32S06), (32S08-32S11), (32S13-32S16),
(32S18-32S21), (32S23-32S26), (32S28-32S31);

---

(3) Phenomenon of Chaotic System Biased Relative Rank Inversion (CSBRRI):

(a) Small decrease in GPA is associated with a large increase in SAPI:
(32S05-32S06), (32S10-32S11), (32S15-32S16),
(32S20-32S21), (32S25-32S26), (32S30-32S31);

(b) Large decrease in GPA is associated a small increase in SAPI:
(32S02>32S03-32S04), (32S07>32S08-32S09),
(32S12>32S13-32S14), (32S17>32S18-32S19),
(32S22>32S23-32S24), (32S27>32S28-32S29);

(c) An even more drastic and shocking illustration of this phenomenon of *Chaotic System Biased Relative Rank Inversion* (CSBRRI): Observe that sometimes –
6 AA's is in fact worse than 1 AA's & 5 AB's : (32S02 < 32S04);
6 AB's is in fact worse than 1 AB's & 5 BB's : (32S07 < 32S09);
6 BB's is in fact worse than 1 BB's & 5 BC's : (32S12 < 32S14);
6 BC's is in fact worse than 1 BC's & 5 CC's : (32S17 < 32S19);
6 CC's is in fact worse than 1 CC's & 5 CD's : (32S22 < 32S24);
6 CD's is in fact worse than 1 CD's & 5 DD's : (32S27 < 32S29);
which correspond to the extreme instances of unfairness in the ubiquitous letter-grading-system.

Table-1.05: Three types of system-intrinsic chaotic behavior (10-point grading scale)

(1) Phenomenon of Chaotic System Biased Amplification or Attenuation of Differentials (CSBAAD):

(a) Small difference in GPA associated with large difference in SAPI:
(22S01-22S03), (22S04-22S05), (22S06-22S08), (22S09-22S10), (22S11-22S13), (22S14-22S15), (22S16-22S18), (22S19-22S20);

(b) Large difference in GPA associated with small difference in SAPI:
(22S02-22S06), (22S07-22S11), (22S12-22S16), (22S17-22S21);

(2) Phenomenon of Chaotic System Biased Suppression or Expression of Differentials (CSBSED):

(a) Large difference in SAPI associated with no difference in GPA:
(22S01-22S02), (22S06-22S07), (22S11-22S12),
(22S16-22S17), (22S21-22S22);

(b) No difference in SAPI associated with large difference in GPA:
(22S03-22S06), (22S08-22S11), (22S13-22S16), (22S18-22S21);

(3) Phenomenon of Chaotic System Biased Relative Rank Inversion (CSBRRI):

(a) Small decrease in GPA is associated with a large increase in SAPI:
(22S05-22S06), (22S10-22S11), (22S15-22S16), (22S20-22S21);

(b) Large decrease in GPA is associated a small increase in SAPI:
(22S02>22S03-22S04), (22S07>22S08-22S09),
(22S12>22S13-22S14), (22S17>22S18-22S19);

(c) An even more drastic and shocking illustration of this phenomenon of *Chaotic System Biased Relative Rank Inversion* (CSBRRI) : Observe that sometimes –

6 A's should have been worse than 1 A's & 5 B's : (22S02 < 22S04);
6 B's should have been worse than 1 B's & 5 C's : (22S07 < 22S09);
6 C's should have been worse than 1 C's & 5 D's : (22S12 < 22S14);
6 D's should have been worse than 1 D's & 5 E's : (22S17 < 22S19);

which correspond to the extreme instances of unfairness in the ubiquitous letter-grading-system.

Table-1.06: Three types of system-intrinsic chaotic behavior
(5-point grading scale)

As another example, consider another group of 38 students, with their academic performance levels (top performers in a class) being very close to one another, identified as 38S01, 38S02, . . . 38S38, each taking the six courses, of which three (s1/4c,s2/4c,s3/4c) are of 4 CCU, two (s4/3c,s5/3c) are of 3 CCU, and one (s6/2c) is of 2 CCU. Suppose that at the end of their academic term, their Raw Scores and Letter Grades in each of the courses are as listed in Table-1.07 [Table-1.08 in case of 5-point scale]. An associated plot of sGPCCU versus sRSCCU is shown in Figure-1.3 [Figure-1.4 in case of 5-point scale]. Again it can be observed that the same chaotic system behavior is exhibited even in this situation. Although the variation in the GPA is only very minor, the corresponding variations in the SAPI can indeed be quite substantial. Also, the more drastic and shocking illustration of the phenomenon of *Chaotic System Biased Relative Rank Inversion* (CSBRRI), can be observed, as presented in Table-1.09 [Table-1.10 in case of 5-point scale].

It is to be noted here that the above observed chaotic system behavior is certainly not because of the specific choice of the set of various parameters and their values, but in fact an intrinsic characteristic of the system, which happen to be exposed very well through these typical example scenarios.

Student ID	Course No. / Credit Units Raw-Score | Letter-Grade												sGPCCU	GPA	GPA-Rank	sRSCCU	SAPI	SAPI-Rank
	s1 / 4c		s2 / 4c		s3 / 4c		s4 / 3c		s5 / 3c		s6 / 2c							
	RS	LG	RS	LG	RS	LG	RS	LG	RS	LG	RS	LG						
38S01	99	AA	99	AA	99	AA	99	AA	99	AA	99	AA	200	10	1	1980	99	1
38S02	90	AA	90	AA	90	AA	90	AA	90	AA	90	AA	200	10	1	1800	90	18
38S03	99	AA	99	AA	99	AA	99	AA	99	AA	89	AB	198	9.9	3	1960	98	2
38S04	90	AA	90	AA	90	AA	90	AA	90	AA	80	AB	198	9.9	3	1780	89	20
38S05	99	AA	99	AA	99	AA	99	AA	89	AB	99	AA	197	9.85	5	1950	97.5	3
38S06	90	AA	90	AA	90	AA	90	AA	80	AB	90	AA	197	9.85	5	1770	88.5	22
38S07	99	AA	99	AA	89	AB	99	AA	99	AA	99	AA	196	9.8	7	1940	97	4
38S08	90	AA	90	AA	80	AB	90	AA	90	AA	90	AA	196	9.8	7	1760	88	23
38S09	99	AA	99	AA	99	AA	99	AA	89	AB	89	AB	195	9.75	9	1930	96.5	5
38S10	90	AA	90	AA	90	AA	90	AA	80	AB	80	AB	195	9.75	9	1750	87.5	24
38S11	99	AA	99	AA	89	AB	99	AA	99	AA	89	AB	194	9.7	11	1920	96	6
38S12	90	AA	90	AA	80	AB	90	AA	90	AA	80	AB	194	9.7	11	1740	87	25
38S13	89	AB	99	AA	99	AA	99	AA	89	AB	99	AA	193	9.65	13	1910	95.5	7
38S14	80	AB	90	AA	90	AA	90	AA	80	AB	90	AA	193	9.65	13	1730	86.5	26
38S15	99	AA	99	AA	99	AA	89	AB	89	AB	89	AB	192	9.6	15	1900	95	8
38S16	90	AA	90	AA	90	AA	80	AB	80	AB	80	AB	192	9.6	15	1720	86	27
38S17	99	AA	99	AA	89	AB	99	AA	89	AB	89	AB	191	9.55	17	1890	94.5	9
38S18	90	AA	90	AA	80	AB	90	AA	80	AB	80	AB	191	9.55	17	1710	85.5	28
38S19	99	AA	99	AA	89	AB	89	AB	89	AB	99	AA	190	9.5	19	1880	94	10
38S20	90	AA	90	AA	80	AB	80	AB	80	AB	90	AA	190	9.5	19	1700	85	29
38S21	99	AA	89	AB	89	AB	89	AB	99	AA	99	AA	189	9.45	21	1870	93.5	11
38S22	90	AA	80	AB	80	AB	80	AB	90	AA	90	AA	189	9.45	21	1690	84.5	30
38S23	89	AB	89	AB	89	AB	99	AA	99	AA	99	AA	188	9.4	23	1860	93	12
38S24	80	AB	80	AB	80	AB	90	AA	90	AA	90	AA	188	9.4	23	1680	84	31
38S25	99	AA	89	AB	89	AB	99	AA	89	AB	89	AB	187	9.35	25	1850	92.5	13
38S26	90	AA	80	AB	80	AB	90	AA	80	AB	80	AB	187	9.35	25	1670	83.5	32
38S27	89	AB	89	AB	89	AB	99	AA	99	AA	89	AB	186	9.3	27	1840	92	14
38S28	80	AB	80	AB	80	AB	90	AA	90	AA	80	AB	186	9.3	27	1660	83	33
38S29	89	AB	89	AB	89	AB	89	AB	99	AA	99	AA	185	9.25	29	1830	91.5	15
38S30	80	AB	80	AB	80	AB	80	AB	90	AA	90	AA	185	9.25	29	1650	82.5	34
38S31	99	AA	89	AB	89	AB	89	AB	89	AB	89	AB	184	9.2	31	1820	91	16
38S32	90	AA	80	AB	80	AB	80	AB	80	AB	80	AB	184	9.2	31	1640	82	35
38S33	89	AB	89	AB	89	AB	99	AA	89	AB	89	AB	183	9.15	33	1810	90.5	17
38S34	80	AB	80	AB	80	AB	90	AA	80	AB	80	AB	183	9.15	33	1630	81.5	36
38S35	89	AB	89	AB	89	AB	89	AB	89	AB	99	AA	182	9.1	35	1800	90	18
38S36	80	AB	80	AB	80	AB	80	AB	80	AB	90	AA	182	9.1	35	1620	81	37
38S37	89	AB	89	AB	89	AB	89	AB	89	AB	89	AB	180	9	37	1780	89	20
38S38	80	AB	80	AB	80	AB	80	AB	80	AB	80	AB	180	9	37	1600	80	38

Table-1.07: A group of 38 students in the top performance band
(10-point grading scale)

Student ID	Course No. / Credit Units Raw-Score | Letter-Grade												sGPCCU	GPA	GPA-Rank	sRSCCU	SAPI	SAPI-Rank
	s1 / 4c		s2 / 4c		s3 / 4c		s4 / 3c		s5 / 3c		s6 / 2c							
	RS	LG	RS	LG	RS	LG	RS	LG	RS	LG	RS	LG						
38S01	99	A	99	A	99	A	99	A	99	A	99	A	100	5	1	1980	99	1
38S02	90	A	90	A	90	A	90	A	90	A	90	A	100	5	1	1800	90	18
38S03	99	A	99	A	99	A	99	A	99	A	89	B	98	4.9	3	1960	98	2
38S04	90	A	90	A	90	A	90	A	90	A	80	B	98	4.9	3	1780	89	20
38S05	99	A	99	A	99	A	99	A	89	B	99	A	97	4.85	5	1950	97.5	3
38S06	90	A	90	A	90	A	90	A	80	B	90	A	97	4.85	5	1770	88.5	22
38S07	99	A	99	A	89	B	99	A	99	A	99	A	96	4.8	7	1940	97	4
38S08	90	A	90	A	80	B	90	A	90	A	90	A	96	4.8	7	1760	88	23
38S09	99	A	99	A	99	A	99	A	89	B	89	B	95	4.75	9	1930	96.5	5
38S10	90	A	90	A	90	A	90	A	80	B	80	B	95	4.75	9	1750	87.5	24
38S11	99	A	99	A	89	B	99	A	99	A	89	B	94	4.7	11	1920	96	6
38S12	90	A	90	A	80	B	90	A	90	A	80	B	94	4.7	11	1740	87	25
38S13	89	B	99	A	99	A	99	A	89	B	99	A	93	4.65	13	1910	95.5	7
38S14	80	B	90	A	90	A	90	A	80	B	90	A	93	4.65	13	1730	86.5	26
38S15	99	A	99	A	99	A	89	B	89	B	89	B	92	4.6	15	1900	95	8
38S16	90	A	90	A	90	A	80	B	80	B	80	B	92	4.6	15	1720	86	27
38S17	99	A	99	A	89	B	99	A	89	B	89	B	91	4.55	17	1890	94.5	9
38S18	90	A	90	A	80	B	90	A	80	B	80	B	91	4.55	17	1710	85.5	28
38S19	99	A	99	A	89	B	89	B	89	B	99	A	90	4.5	19	1880	94	10
38S20	90	A	90	A	80	B	80	B	80	B	90	A	90	4.5	19	1700	85	29
38S21	99	A	89	B	89	B	89	B	99	A	99	A	89	4.45	21	1870	93.5	11
38S22	90	A	80	B	80	B	80	B	90	A	90	A	89	4.45	21	1690	84.5	30
38S23	89	B	89	B	89	B	99	A	99	A	99	A	88	4.4	23	1860	93	12
38S24	80	B	80	B	80	B	90	A	90	A	90	A	88	4.4	23	1680	84	31
38S25	99	A	89	B	89	B	99	A	89	B	89	B	87	4.35	25	1850	92.5	13
38S26	90	A	80	B	80	B	90	A	80	B	80	B	87	4.35	25	1670	83.5	32
38S27	89	B	89	B	89	B	99	A	99	A	89	B	86	4.3	27	1840	92	14
38S28	80	B	80	B	80	B	90	A	90	A	80	B	86	4.3	27	1660	83	33
38S29	89	B	89	B	89	B	89	B	99	A	99	A	85	4.25	29	1830	91.5	15
38S30	80	B	80	B	80	B	80	B	90	A	90	A	85	4.25	29	1650	82.5	34
38S31	99	A	89	B	89	B	89	B	89	B	89	B	84	4.2	31	1820	91	16
38S32	90	A	80	B	80	B	80	B	80	B	80	B	84	4.2	31	1640	82	35
38S33	89	B	89	B	89	B	99	A	89	B	89	B	83	4.15	33	1810	90.5	17
38S34	80	B	80	B	80	B	90	A	80	B	80	B	83	4.15	33	1630	81.5	36
38S35	89	B	89	B	89	B	89	B	89	B	99	A	82	4.1	35	1800	90	18
38S36	80	B	80	B	80	B	80	B	80	B	90	A	82	4.1	35	1620	81	37
38S37	89	B	89	B	89	B	89	B	89	B	89	B	80	4	37	1780	89	20
38S38	80	B	80	B	80	B	80	B	80	B	80	B	80	4	37	1600	80	38

Table-1.08: A group of 38 students in the top performance band
(5-point grading scale)

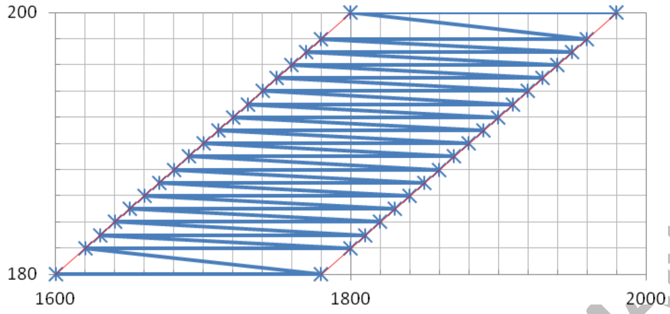

Figure-1.3: Plot of sGPCCU versus sRSCCU for 38 students
(10-point grading scale)

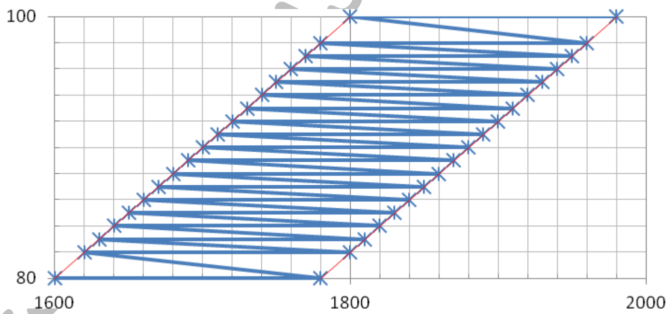

Figure-1.4: Plot of sGPCCU versus sRSCCU for 38 students
(5-point grading scale)

6 AA's & 0 AB's should have been worse than 5 AA's & 1 AB's:
38S02 < 38S03; 38S02 < 38S05; 38S02 < 38S07;
6 AA's & 0 AB's should have been worse than 4 AA's & 2 AB's:
38S02 < 38S09; 38S02 < 38S11; 38S02 < 38S13; 38S02 < 38S15;
6 AA's & 0 AB's should have been worse than 3 AA's & 3 AB's:
38S02 < 38S17; 38S02 < 38S19; 38S02 < 38S21; 38S02 < 38S23;
6 AA's & 0 AB's should have been worse than 2 AA's & 4 AB's:
38S02 < 38S25; 38S02 < 38S27; 38S02 < 38S29;
6 AA's & 0 AB's should have been worse than 1 AA's & 5 AB's:
38S02 < 38S31; 38S02 < 38S33;
5 AA's & 1 AB's should have been worse than 4 AA's & 2 AB's:
38S04 < 38S09; 38S04 < 38S11; 38S04 < 38S13; 38S04 < 38S15;
5 AA's & 1 AB's should have been worse than 3 AA's & 3 AB's:
38S04 < 38S17; 38S04 < 38S19; 38S04 < 38S21; 38S04 < 38S23;
5 AA's & 1 AB's should have been worse than 2 AA's & 4 AB's:
38S04 < 38S25; 38S04 < 38S27; 38S04 < 38S29;
5 AA's & 1 AB's should have been worse than 1 AA's & 5 AB's:
38S04 < 38S31; 38S04 < 38S33; 38S04 < 38S35;
4 AA's & 2 AB's should have been worse than 3 AA's & 3 AB's:
38S10 < 38S17; 38S10 < 38S19; 38S10 < 38S21; 38S10 < 38S23;
4 AA's & 2 AB's should have been worse than 2 AA's & 4 AB's:
38S10 < 38S25; 38S10 < 38S27; 38S10 < 38S29;
4 AA's & 2 AB's should have been worse than 1 AA's & 5 AB's:
38S10 < 38S31; 38S10 < 38S33; 38S10 < 38S35;
4 AA's & 2 AB's should have been worse than 0 AA's & 6 AB's:
38S10 < 38S37;
3 AA's & 3 AB's should have been worse than 2 AA's & 4 AB's:
38S16 < 38S25; 38S16 < 38S27; 38S16 < 38S29;
3 AA's & 3 AB's should have been worse than 1 AA's & 5 AB's:
38S16 < 38S31; 38S16 < 38S33; 38S16 < 38S35;
3 AA's & 3 AB's should have been worse than 0 AA's & 6 AB's:
38S16 < 38S37;
2 AA's & 4 AB's should have been worse than 1 AA's & 5 AB's:
38S26 < 38S31; 38S26 < 38S33; 38S26 < 38S35;
2 AA's & 4 AB's should have been worse than 0 AA's & 6 AB's:
38S26 < 38S37;
1 AA's & 5 AB's should have been worse than 0 AA's & 6 AB's:
38S32 < 38S37.

Table-1.09: Phenomenon of *Chaotic System Biased Relative Rank Inversion* (CSBRRI) : GPA comparison - contrary to the teacher's perspective (10-point grading scale)

6 A's & 0 B's should have been worse than 5 A's & 1 B's:
38S02 < 38S03; 38S02 < 38S05; 38S02 < 38S07;
6 A's & 0 B's should have been worse than 4 A's & 2 B's:
38S02 < 38S09; 38S02 < 38S11; 38S02 < 38S13; 38S02 < 38S15;
6 A's & 0 B's should have been worse than 3 A's & 3 B's:
38S02 < 38S17; 38S02 < 38S19; 38S02 < 38S21; 38S02 < 38S23;
6 A's & 0 B's should have been worse than 2 A's & 4 B's:
38S02 < 38S25; 38S02 < 38S27; 38S02 < 38S29;
6 A's & 0 B's should have been worse than 1 A's & 5 B's:
38S02 < 38S31; 38S02 < 38S33;
5 A's & 1 B's should have been worse than 4 A's & 2 B's:
38S04 < 38S09; 38S04 < 38S11; 38S04 < 38S13; 38S04 < 38S15;
5 A's & 1 B's should have been worse than 3 A's & 3 B's:
38S04 < 38S17; 38S04 < 38S19; 38S04 < 38S21; 38S04 < 38S23;
5 A's & 1 B's should have been worse than 2 A's & 4 B's:
38S04 < 38S25; 38S04 < 38S27; 38S04 < 38S29;
5 A's & 1 B's should have been worse than 1 A's & 5 B's:
38S04 < 38S31; 38S04 < 38S33; 38S04 < 38S35;
4 A's & 2 B's should have been worse than 3 A's & 3 B's:
38S10 < 38S17; 38S10 < 38S19; 38S10 < 38S21; 38S10 < 38S23;
4 A's & 2 B's should have been worse than 2 A's & 4 B's:
38S10 < 38S25; 38S10 < 38S27; 38S10 < 38S29;
4 A's & 2 B's should have been worse than 1 A's & 5 B's:
38S10 < 38S31; 38S10 < 38S33; 38S10 < 38S35;
4 A's & 2 B's should have been worse than 0 A's & 6 B's:
38S10 < 38S37;
3 A's & 3 B's should have been worse than 2 A's & 4 B's:
38S16 < 38S25; 38S16 < 38S27; 38S16 < 38S29;
3 A's & 3 B's should have been worse than 1 A's & 5 B's:
38S16 < 38S31; 38S16 < 38S33; 38S16 < 38S35;
3 A's & 3 B's should have been worse than 0 A's & 6 B's:
38S16 < 38S37;
2 A's & 4 B's should have been worse than 1 A's & 5 B's:
38S26 < 38S31; 38S26 < 38S33; 38S26 < 38S35;
2 A's & 4 B's should have been worse than 0 A's & 6 B's:
38S26 < 38S37;
1 A's & 5 B's should have been worse than 0 A's & 6 B's:
38S32 < 38S37.

Table-1.10: Phenomenon of *Chaotic System Biased Relative Rank Inversion* (CSBRRI) : GPA comparison - contrary to the teacher's perspective (5-point grading scale)

The chaotic system behavior is neither arising from nor is avoidable by any teachers through their decisions on the evaluation itself. Also, such system behavior is neither associated with nor can be avoided by any general or specific students' actions through their academic performance levels, whether individually or otherwise. The various parameters, like the number of courses considered for the analysis, the Course Credits/Units (CCU), the number of students being considered, the relative performance levels among them, the actual distribution of the raw scores in a class, and even the actual letter grading system model (policies and procedures) adopted, can in fact possibly be changed within their corresponding ranges of usually acceptable variability, and still we can construct any number of example scenarios like the above, to illustrate such system behavior.

What do we conclude from the above observations? It is true that every teacher really means by whatever s/he initially gives as the raw scores (marks) for each of the test papers. All other transformations, conversions or computations, are only performed later, starting with these raw scores as the original raw input data; in order just to conform to the Policies and Procedures of the prevailing Evaluation System as per the Academic Rules and Regulations in the school/institute. However, the "*Student Academic Performance Indiex* (SAPI)" computed directly as a weighted average of the *original Raw Scores*, is certainly the best and in fact a true representation, of whatever the teachers really intended (collectively) while originally evaluating each of their students.

Unfortunately, the prevailing LGS System just does not provide a reliable mechanism for a teacher to convey the information regarding the student academic performance evaluation, through the transcripts/grade reports, to whoever is really concerned in receiving (and possibly even acting, based on) such information! The system behavior exhibiting the *Phenomenon of Chaotic System Biased Amplification or Attenuation of Differentials* (CSBAAD), the *Phenomenon of Chaotic System Biased Suppression or Expression of Differentials* (CSBSED), and the *Phenomenon of Chaotic System Biased Relative Rank Inversion* (CSBRRI), is certainly extremely undesirable, or rather just simply *unacceptable*.

The system is grossly *unfair to the students*, because they get subjected to *chaotic system biased unfair and unreliable comparisons* (CSBUUC), of course, not caused by any individual person, but because of the poor system design, which again is not a result of any deliberate intent of anyone. The system is *unfair to the recruiters* or prospective employers, because they are *misled by unreliable and/or misleading information*; when in fact, they deserve to be provided with the best and most reliable information on the students academic performance; by unbiased and fair measurements, with appropriately designed reliable performance indicators, and communicated through the transcripts or grade reports. The system is *unfair to the teachers* who although *helpless* (because of the existing system imposed on them, to be followed) are looked upon as the possible perpetrators of such a disturbingly chaotic system, which is both unfair as well as unreliable; wherein the Raw Scores/Marks that they had originally

assigned, are later subjected to chaotic system biased information loss/corruption. Also, the LGS system design wrongly presumes that the teacher's precision in evaluation is rather poor, limited to classification into possibly only a handful of distinct categories. However, there is an intrinsic contradiction in the system design philosophy itself since a significantly higher precision level is mysteriously presumed to have been achieved in the final figures of GPA, CGPA, etc.

1.4 Summary of Observations

The data presented here clearly shows the highly undesirable system behavior resulting in severe unfairness in the associated comparisons among the students; and that too not just among those in a small neighborhood of performance levels, but across the entire spectrum of performance levels. Further detailed analysis of such chaotic system characteristics is presented in the second part - *Academic Performance Evaluation of Students* - U*biquitous System Analyzed* (APES-USA) - clearly indicating that the root cause is some serious lacunae inherent in its very design. The third part presents the "Student Academic Performance Evaluation System (SAPES) - A Proposed Solution to the APES-USA Problem" as an alternative system design.

An Interlude - with the Author

Good to see you here, having read through the first part of the book. There, we have sifted through various scenarios to get ourselves convinced of the fact that the letter grading system does indeed exhibit a highly chaotic behavior, resulting in grossly unfair comparisons across the entire spectrum of possible performance levels of students; that it is neither limited to any narrow band of performance levels, nor limited by any case of some remote probabilistic coincidences.

Let us quickly recapitulate the salient features of our observations:

(1) The (overall aggregate) raw scores/marks (of all the tests/etc) as originally awarded by the teacher(s), is the most reliable measure of a student's academic performance level in the course(s), that teacher(s) can confidently bank upon for fair comparisons.

(2) The use of a grade point scale (for the purpose of awarding the letter grades) as imposed on the teacher by the academic rules & regulations of the school, requires the teacher to perform a mapping from the raw scores/marks onto the grade points.

(3) The conversion of raw scores/marks to grade points (letter grades) - is the *first step* which introduces information loss/corruption.

(4) The computation of the GPA and/or CGPA figures, through a weighted average aggregation process - is the *second step* which introduces information loss/corruption.

(5) Specifically, three kinds of system-intrinsic phenomena of Chaotic System Biased Information Loss/Corruption (CSBILC) have been identified: Chaotic System Biased Amplification or Attenuation of Differentials (CSBAAD), Chaotic System Biased

Suppression or Expression of Differentials (CSBSED), Chaotic System Biased Relative Rank Inversion (CSBRRI).

(6) These system-intrinsic phenomena of chaotic system biased information loss/corruption result in Chaotic System Biased Unfair and Unreliable Comparisons (CSBUUC) among the students being evaluated, thus making the LGS system to be very grossly unfair to the students. It is also unfair to the prospective employers who expect some relevant, unbiased, reliable information to be contained in and conveyed through those transcripts or grade reports. Again, it is very unfair to the teachers, who find themselves utterly helpless, when the raw scores/marks that they had originally awarded very conscientiously, are later subjected to chaotic system biased information loss/corruption.

The most discussed current hot topic in academia is the problem of Grade Inflation! *Grade Inflation is in fact just a symptom, the root cause there-for being the vulnerability* (*non-robustness*) *of the academia to extraneous influences, thus rendering the entire evaluation unreliable! It is a* human issue*, which needs to be dealt with as such, at that level itself,* irrespective of *whatever be the procedural system design for the academic performance evaluation of students.* Therefore, the problem of grade inflation is not dealt with any further.

In the second part, we will go through a thorough and systematic analysis of the LGS system characteristics, to show that such system-intrinsic chaotic behavior is in fact very much inherent as characterized by its very design; the root cause being the use of a very crude grade point scale - with an extremely poor scale precision.

PART - II

An Analysis

APES-USA

Academic Performance Evaluation of Students - Ubiquitous System Analyzed

2.1 Rounding Errors in Aggregation Processes

In order to start an analysis of the typical Letter Grading System (LGS) model as presented in the earlier part of this report, let us first pose a simple problem to ourselves and try to find an acceptable solution. Let us look at a scenario wherein you are required to measure the length of the perimeter of a pentagon, and report the result in meters (say, p), and then also to compute the average length of a side of that pentagon and again report that in meters (say, s). However, suppose we first measure the length of each of the five sides of the pentagon in meters, and then round-off each of these figures to the nearest meter, add them up, and report the result (say, _Rp), and using this we compute the average length of a side in meters (say, _Rs). What would be the error in such result?

Let us say, for the sake of concreteness, that there are four pentagons P1, P2, P3, P4, with the actual lengths of their sides (in m), the perimeter p and the average length of a side s, being as follows:

[see Table-2.1]

P1: {1.8, 1.7, 1.6, 1.2, 1.1}; p1 = 7.4 m, and s1 = 1.48 m;
P2: {1.9, 1.8, 1.4, 1.3, 1.2}; p2 = 7.6 m, and s2 = 1.52 m;
P3: {1.7, 1.6, 2.3, 2.2, 2.1}; p3 = 9.9 m, and s3 = 1.98 m;
P4: {1.9, 1.8, 1.7, 2.4, 2.3}; p4 = 10.1 m, and s4 = 2.02 m.

If we were to use the *standard round-off* for the length measurement of every side to nearest meter, the results would be:

P1: {2, 2, 2, 1, 1}; sRp1 = 8 m, and sRs1 = 1.6 m;
P2: {2, 2, 1, 1, 1}; sRp2 = 7 m, and sRs2 = 1.4 m;
P3: {2, 2, 2, 2, 2}; sRp3 = 10 m, and sRs3 = 2.0 m;
P4: {2, 2, 2, 2, 2}; sRp4 = 10 m, and sRs4 = 2.0 m.

If we were to use the *rounding-down* to the Floor for the length measurement of every side, the results would be:

P1: {1, 1, 1, 1, 1}, rDp1 = 5 m, and rDs1 = 1.0 m;
P2: {1, 1, 1, 1, 1}, rDp2 = 5 m, and rDs2 = 1.0 m;
P3: {1, 1, 2, 2, 2}, rDp3 = 8 m, and rDs3 = 1.6 m;
P4: {1, 1, 1, 2, 2}, rDp4 = 7 m, and rDs4 = 1.4 m.

If we were to use the *rounding-up* to the Ceiling for the length measurement of every side, the results would be:

P1: {2, 2, 2, 2, 2}, rUp1 = 10 m, and rUs1 = 2.0 m;
P2: {2, 2, 2, 2, 2}, rUp2 = 10 m, and rUs2 = 2.0 m;
P3: {2, 2, 3, 3, 3}, rUp3 = 13 m, and rUs3 = 2.6 m;
P4: {2, 2, 2, 3, 3}, rUp4 = 12 m, and rUs4 = 2.4 m.

Now, we notice that p1 < p2; p3 < p4 and thus s1 < s2; s3 < s4. The standard rounding process introduces error to the extent that the resultant aggregates have their relative values as though reversed in one case and suppressed in the other case; that is, sRp1 > sRp2 and sRp3=sRp4. On the other hand, the rounding down and rounding up processes seem to have introduced errors so as to cause the suppression of the differences in the first case and a reversal of the differences in the second case; that is, rDp1 = rDp2 and rUp1 = rUp2 whereas rDp3 > rDp4 and rUp3 > rUp4.

P1	roundDown(P1)	Round(P1)	roundUp(P1)
1.8	1	2	2
1.7	1	2	2
1.6	1	2	2
1.2	1	1	2
1.1	1	1	2
7.4	5	8	10
P2	roundDown(P2)	Round(P2)	roundUp(P2)
1.9	1	2	2
1.8	1	2	2
1.4	1	1	2
1.3	1	1	2
1.2	1	1	2
7.6	5	7	10
P3	roundDown(P3)	Round(P3)	roundUp(P3)
1.7	1	2	2
1.6	1	2	2
2.3	2	2	3
2.2	2	2	3
2.1	2	2	3
9.9	8	10	13
P4	roundDown(P4)	Round(P4)	roundUp(P4)
1.9	1	2	2
1.8	1	2	2
1.7	1	2	2
2.4	2	2	3
2.3	2	2	3
10.1	7	10	12

Table-2.1: Rounding Errors in Aggregation.

The results associated with the above round-off procedure certainly suffer from *Information Loss/Corruption*; because of the *Round-off Errors*. Such a system of measurement and/or computation, exhibits a chaotic, highly complex, counter-intuitive as well as undesirable system behavior, due to the Sharp Interval Domain Boundaries across which distinctly different decisions are followed regarding the round-off procedure.

One may remind oneself of the well established standard procedure usually adopted in numerical computations to use double precision arithmetic for all intermediate computations, although the input as well as the expected output may happen to be in single precision representations. The double precision representation of all the intermediate computations allows sufficient room to accommodate for the cumulative errors of computations (arising due to the finite word size arithmetic used in computers) so that at the end the result would still be acceptable with a single precision accuracy level. On the other hand, imagine (as is well known) what would be the consequences (in terms of the errors involved) if the input and output were expected/required to be in double precision, whereas all the intermediate computations were to be performed in single precision arithmetic.

What does this scenario have to do with the letter grading system; namely the conversion of the raw scores to letter grades, and then to the grade points, and finally the computation of the GPA and/or CGPA as the weighted average of these Grade Points? Yes,

indeed there is more than just a semblance of an analogy here! The sides of the pentagon could be the five courses, their lengths would be the raw scores the perimeter would be the total score (p, the total raw score; or sRp, the sum total of the product of the grade points for each courses multiplied by the corresponding CCU, as the numerator computed in the GPA calculation); and the computed average side length would be the overall average score (s, the overall average score computed directly as a weighted average of the raw scores; or sRs the usual GPA).

The situation in the above example would be same irrespective of which specific round-off procedure is used, like rounding downward to the Floor, or rounding upward to the Ceiling, instead of the standard round-off procedure. A further look into the details of the system behavior shows without any ambiguities, that such a system behavior arises not particularly from the specific mechanism or procedure for the round-off, but rather from the fact that there has been a round-off errors introduced by the procedure applied in the measurement and/or computational system.

2.2 Quantization, Contraction-Expansion Mappings, Information Loss/Corruption

Before we take up the task of proposing an alternative system design, we need to first probe deeper into the causes for the above observed system behavior. The measurement scale for raw score is

usually of a better (smaller) precision level than that of the *Grade Points Scale* used for the letter grades. *The Scale Precision is generally defined as a fraction of* (*relative to*) *the Scale Range*. However, in the case of *quantized discrete scales* (with a discrete set of points, as the only valid measurements possible, rather than having an interval range) the scale-precision may equivalently be expressed as the reciprocal of the number that is *one less than the cardinality* of the set of valid quantized discrete measurement values that such a *quantized discrete scale* accepts. The set of possible raw score values that a teacher may use in actual evaluation, say, using a scale range of 0-100 and a precision of 0.01 or 1-in-100, thus giving rise to 101 distinctly different possible values for a raw score, or equivalently, 101 distinct points as markings on the raw scores scale, labeled by the integers from 0 to 100, is of cardinality 101. However, the set of acceptable or valid letter grades (each associated with a grade point value used for computation of GPA and/or CGPA) is usually of a small cardinality; for example, at most 5, or 6, or 11, or 16.

Any mapping scheme that one designs, to convert *Raw Score* to *Letter Grade*, would necessarily be a *Contraction Mapping Scheme*; that is, the "range" of the mapping is of smaller size than the "domain" of the mapping. Further, any contraction mapping between discrete sets, is necessarily a many-to-one mapping, thus resulting in *Information Loss/Corruption*.

Even supposing that we are considering the measurement scale to be continuous within its range, usually the mapping schemes

that are prevalent, as is the one in the typical LGS system model explained above, do incorporate some *Quantization* of the raw scores scale. In other words, the raw scores scale is divided into some appropriate fixed number of *Quantization Interval Domains*, each of which is *mapped onto a single point/value in the discrete scale* of the set of letter grades; thus resulting in an effective *non-selectivity* or *non-differentiability* within each of these quantization interval domains. The fact that, for later computations, each letter grade is associated with a point/value on *a seemingly continuous scale of the Grade Points*, is yet another mysterious system characteristic feature, that seems to have been incorporated in the overall system design. Anyway, it is essential to note here, that every instance of quantization is invariably associated with the occurrence of sharp *Quantization Interval Domain Boundaries*, which are points of *discontinuities* in the system behavior.

It is to be noted that in any general Information System, when one passes through a transformation among the several alternative information representation schemes, there is a possibility of *Information Loss/Corruption* associated with such instances of *Quantization* and/or *Contraction Mapping*, followed by an implicit *Expansion Mapping* (corresponding to the mysterious presumption of a significantly higher degree of precision in the resultant combined aggregate value; the GPA and/or the CGPA here). The only case that would not result in Information Loss/Corruption, is the one with the *Raw Scores Scale* and the *Grade Points Scale* to have exactly the same Range and Precision, thus having a *one-to-one onto mapping*

transformation, for conversion between them. But, of course, this is far from reality, as observed in actual practice. From the point of view of Computational Systems Theory, the above observed *Information Loss/Corruption*, as well as the *chaotic and counter-intuitive system behavior*, are caused by the *drastically poor scale precision* used in the *intermediate representations*, that is, by the set of letter grades, or equivalently, the Grade Points Scale.

In the above example, the scale precision associated with the raw scores scale is 0.01 or 1-in-100. All the intermediate computations are performed on a very crude/rough/coarse scale of very poor precision; associated with each of the several discrete decisions regarding the specific choice of the letter grade to be assigned to each of the students in each of the subjects, from among the available set of assignable letter grades, that is of cardinality, say 6, thus having a scale precision of 0.20 or 1-in-5. However, the precision expected at the end, in the GPA and/or the CGPA as reported in (or, otherwise invariably determined, from) the transcripts/grade-reports is about 0.002 or 0.01-in-5.00 (that is, an accuracy of upto the second decimal place in the GPA and/or the CGPA figure, on a scale range of 0-5). This significantly higher precision level somehow mysteriously presumed to have been achieved in the final figures of GPA and/or CGPA etc, is implied from the Expansion Mapping used in their computations as the weighted average of the Grade Points corresponding to the letter grades. This is exactly like the case of using single precision arithmetic for all intermediate computations wherein the initial input

data is in double precision and the final result is also expected to be in double precision. The lost information cannot be recovered, but only get corrupted by this computational approach! Now, imagine how it is possible at all, to expect *unbiased, fair, and reliable* representation and transfer of information by/through such a system!

Also, it is to be noted here that the use of Course Credits/Units (CCU) as multiplicative weighting factors to the Grade-Points (as well as the Grade-Points-Scale range) associated with the letter-grades, cannot result in any improvement in the scale precision, although the range does get widened/enlarged, since the scale quantum interval domain size also gets widened/enlarged by exactly the very same multiplicative factor.

2.3 Measurement-Scale: Type, Range, Precision, and Scale Fixed Points

Any rational or systematic evaluation is possible only through appropriate measurements of the characteristic properties or criteria for such evaluation. Any measurement and therefore any evaluation is possible only through comparisons; and even the very purpose of evaluation is for comparison among the entities evaluated using the criteria being measured. If it were not for comparisons, no measurement is possible, nor any evaluation required/possible.

Any measurement system must incorporate an appropriately designed scale for the measurements. The nature of the entities being characterized through such measurements, and also the kind of the specific attribute that is being measured, determine the type of the measurement scale that may be appropriate for the purpose. These details lead to the establishment of some well-accepted standards for such measurements, with appropriate measurement scales defined. These measurement standards and the associated scales for measurements may be classified into several broad categories.

First consider a *nominal scale*, that is used for characterization of nominal entities that require some kind of distinct labeling of the different possible attribute values, with no specific relational order among them. For example, the color of some object may be red whereas that of another may be blue, and this attribute (maybe along with other possible attributes, in combination) may be used in the characterization and/or the distinct identification of one object from another. So, a finite *discrete set* of colors from a possible collection of identifiable colors in a palette (based on the technology available) may serve the purpose.

Very soon we end up in a situation which forces us to improve our measurement system, if we think of two or more objects that are somewhat gray, somewhere between pure white and pure black. One may extend the finite discrete set of colors in the palette, but that process may result in a very unwieldy measurement scale to work with in practice. So, a reasonable relational operator is

imposed on the attributes, so that there is a structure for the set of valid values that an attribute can take. In the case of colors, there exists a well established system to characterize the visual appearance of an object from the point of view of the measurements associated with the color, hue, brightness, etc. of the light received from the object (reflected or deflected or emitted, etc) described through appropriate parameter values (e.g., the RGB co-ordinates and the intensity/brightness, or even a detailed Spectral Distribution data). A measurement scale when used for relative comparisons among values that an attribute takes on the scale, corresponding to two or more objects, can be considered to be a *relational scale*. For example, one object may be of relatively lighter blue color, when compared with a second object that is of darker blue color, and yet another object that is of deep blue color. Here, it may not be meaningful to take the numerical differences in the attribute values, and say for example, that the first object is as much lighter in its blue color when compared to the second object, as the third object is darker in its blue color when compared with the second object. In other words, although the measurement scale admits relational operators (comparisons), the numerical differences, or the interval sizes, or the distances, among the attribute values may not be meaningfully defined!

Now consider, as an example, the Centigrade (or even a Fahrenheit) scale for temperature measurement, using which the temperatures of two or more objects can be compared, in the sense than one could be either more or less than another by some

extent/degree, and such differences in extent/degree can be well characterized using the same measurement scale. A measurement scale that allows the direct measurement of (or a later computation) of the extent/degree of the differences in attribute values (through some well defined concept of interval size, or that of a distance measure, or simply the numerical difference) can be considered to be a *relative interval scale*.

In an interval scale, although the interval sizes, or the distances, or simply the numerical differences in the attribute values, are well defined, there may not exist any absolute fixed point in such a measurement scale. Now, if instead of the Centigrade or the Fahrenheit scales, we consider the Absolute Temperature Scale (degree-Kelvin) it turns out that there exists what can be considered as an *Absolute Fixed Point* in such a scale. The very measurement in such a scale is a direct and explicit comparison of an attribute value relative to that absolute fixed point. Such a measurement scale, with some (one or sometimes more) absolute fixed points, is considered as an *absolute interval scale*. The attribute values measured in an absolute interval scale may allow for (1) meaningful relative comparisons among attribute values, (2) meaningful concept of the interval sizes, or a distance measure, or simply the numerical differences in the attribute values, (3) meaningful computations of the numerical ratios of the interval sizes, (4) meaningful concepts associated with the distances, w.r.t the absolute fixed point/points, and their ratios.

2.4 Scale Quantum-Unit Interval-Domain Size (SQUIDS)

A complete description of a measurement scale must necessarily include the specification of the *Scale Range* (*cardinality of the set of possible values, in the case of a quantized finite discrete scale*) and the *Scale Precision or the Scale Quantum Unit Interval Domain Size* (SQUIDS), and also the *Scale Fixed Points*, if any.

At this point, it is very useful to note how the incorporation of an appropriately chosen design modification in a system, although seemingly simplistic, can in fact result in a significant enhancement in the system characteristics. Specifically, consider the incorporation of a negative one-fourth (-0.25) mark/score for every wrong answer, in a written aptitude test with 100 questions and each correct answer carrying one ($+1$) mark/score. It is very easy to see that the scale range gets enlarged/widened; that is, from -25 to $+100$ now, as against the earlier range namely, from 0 to 100. The scale quantum unit interval domain size gets reduced to 0.25 now, as against 1.00 earlier. As a result, there is a five-fold improvement in the scale precision, that is, 0.002 or 0.25-in-125 or 1-in-500 now, as against 0.01 or 1-in-100 earlier.

A finer scale, with a better precision allows the teacher to incorporate finer distinctions in the various possible performance levels encountered during the evaluation process, and does not force the teacher to be restricted by the limitations of a

coarse/crude/rough scale with poorer precision. A coarser scale (poorer precision) gives rise to undue magnification of the differences between the possible measurements that happen to be located on either side just across the border between adjacent scale quantum unit interval domains. Of course, if a teacher does not need the precision provided by the measurement scale incorporated in the system design; a suitable mapping or superimposition of an appropriately chosen coarser scale onto the finer scale (incorporated in the system design) can in fact be performed with neither any Information Loss/Corruption nor any undesirable system behavior like the ones observed and analyzed here. For example, a teacher may decide to use only the eleven numbers 0, 10, 20, . . . 100; instead of the 101 possibilities, even though provided with a percentage scale with the scale range 0-100 and scale precision of 1-in-100. It is generally advisable however, for each teacher to effectively utilize the full representational capability, and also the associated selectivity (or resolution, or grain size or precision) that is made available through the measurement system design.

From the analyses presented above, it is clear that the most fundamental basis in the design of any measurement scale is in fact, the *Scale Quantum Unit Interval Domain Size* (SQUIDS), *which defines the absolute Scale Precision*, which together with the *Scale Range* and the *Scale Fixed Points (if any)* provides a complete description for the measurement system.

From the point of view of quantization, the Raw Scores Scale can also be considered as a quantized finite discrete scale, although with a smaller scale quantum unit interval domain size. A contraction mapping from the raw scores scale to the *Grade Points Scale* (set of letter grades) necessarily results in the an effective lumping together of some of the neighboring quantum unit interval domains into an agglomeration, thus resulting in an enlarged/widened the effective Scale Quantum Unit Interval Domain Size when expressed as a fraction of the corresponding scale range. So, the effect of contraction mapping in the case of quantized finite discrete scales, is exactly the same as that of quantization on continuous scales; and that is, not only to worsen the scale precision, but also to introduce *discontinuities in the system characteristics*. This is how a situation arises wherein not only some crucial information is lost, but also a manifestation of the chaotic, highly complex, counter-intuitive, seemingly unpredictable system behavior, is observed, effectively introducing chaotic system biased information corruption as well.

2.5 Diagnostic Analysis of the LGS chaotic behavior

The undesirable system behavior explained earlier, is the result of a poor system design, specifically in terms of the implicit changes both in the *type of the scale* as well as the *scale range* and particularly the *scale precision* or the *scale quantum unit interval domain size*, as incorporated into the system design at the different

stages of information flow through the system. Or, rather there is a *possible ambiguity* as to the appropriate application of the very concept of scale quantum unit interval domain size, and that of scale precision (only the concept of scale range seems to have been explicitly incorporated) even if implicit in the system design.

Specifically, in the first stage the teacher is given an almost complete freedom as to the *definition of a measurement scale* to be used for raw scores (marks); the system design does not address this issue at all. The system design is ambiguous in the second stage, wherein a mapping scheme or transformation is to be defined to convert the raw scores to the letter grades defined as a finite discrete set of relatively small cardinality (on which is possibly defined an anti-symmetric transitive binary relation, "better-than"). This results in a drastic deterioration of the scale precision (defined as the scale quantum unit interval domain size). In the third stage, there is an almost immediate usage of a well defined one-to-one onto mapping to the *Grade Points Scale*, as defined and specified by the system design, thus maintaining the same very low/poor precision. In the fourth stage, the Course Credits/Units (CCU) are used as multiplicative weighting factors in the computation of a combined aggregate or overall score (namely, the GPA or the CGPA) on the *Grade Points Scale*, with an mysterious expectation of an impossibly high precision level as mentioned earlier.

In addition to these changes in the nature of measurement scale used, the use of the Course Credits/Units (CCU) as

multiplicative weighting factors in the fourth stage, that is the computation of the overall score (namely, the GPA) is again not a good systems design. For example, consider a situation wherein a teacher evaluates two students, the first one being at 90% level, and the second one being at 89% level, using a percentage scale, their difference being just equal to the scale quantum unit interval domain size (SQUIDS). When multiplied by the CCU value of 4, for example, these numbers become 360 and 356 respectively. Now, let us ask the question, why shouldn't the teacher be allowed to use the enhanced scale range directly for the purpose of the evaluation? It might so happen that, if the teacher were to be given this facility of an enhanced scale range right at the initial stage of evaluation itself, with the same fixed value of the scale quantum unit interval domain size (SQUIDS), that is, the absolute scale precision, then, the actual evaluations (raw score or marks) could possibly as well have been anywhere in the range 358-361 for the first student, and in the range 354-357 for the second student, out of 400 points. This design feature would have given a better measurement scale on the hands of the teacher right at the initial step of evaluation, and thus improve the effective resolution (or selectivity or grain-size or precision) of the entire evaluation system. From this point of view, the incorporation of (Course Credit Units as the) multiplicative weighting factors in the computation of the overall aggregate score, or the GPA and/or the CGPA, is in fact *analogous to a delayed patching up for a lost opportunity without any real benefit*.

It is to be noted here that, in situations where numerical multiplication and/or division operations are carried out, the concept of *relative precision* expressed as a fraction of (relative to) the scale range, is relevant; whereas, in situations where numerical addition and/or subtraction operations are carried out, the concept of *absolute precision* is relevant; both being defined in terms of the scale quantum unit interval domain size (SQUIDS).

2.6 A Possible Mathematical System Model

Although a more rigorous mathematical system model is beyond the scope of this book, a brief outline of a possible approach towards such a general mathematical model for a generic letter grading system is presented here below.

The sequence of intervals of marks (raw scores) corresponding to each of the grades (grade points) is defined usually by fixing the lower and upper cut-off marks on the raw scores scale corresponding to each of the grade points. This may be defined for each of the courses by the respective teacher; either as a fixed mapping, in a typical distribution independent grading scheme; or depending on the actual distribution of the overall raw scores of all the students in the class, in a typical distribution based grading scheme, irrespective of whether an attempt is made to force-fit the resultant grades onto some standard normalized distribution curve.

Such a sequence of interval domains in the raw scores scale corresponding to each of the grades (grade points) on the grade points scale, as defined by the teacher for the specific course; maybe represented by the corresponding lower cut-off marks m(c, g) and upper cut-off marks n(c, g). In other words, the entire raw scores scale is partitioned into *course specific partition* of non-overlapping interval domains [each interval therein in turn being defined by its lower and upper cut-off marks] as defined by the teacher.

The *computation* of GPA and/or CGPA can be considered as a *convex combination of the corresponding grade point values*, with the relative course credit units as the weighting factors. The mapping of that *computation* onto the corresponding intervals in the raw scores scale indicates the possible variations in the SAPI values that would still result in the same given GPA and/or CGPA figures. Because each grade point gets mapped onto an interval domain in the raw scores scale, the GPA (and/or CGPA) that is a convex combination of the grade point values get mapped onto a convex combination of the corresponding interval domains on the raw scores scale. A two dimensional representation of this computational process produces a convex hull of the end points of all those interval domains corresponding to all those grade point values.

For example, suppose there are six courses, then the convex hull generated will be from the end points of the marks-interval-domains for each of these six course, comprising of the twelve points on a plot of grade points versus marks - (i) l(c1,g1); (ii) u(c1,g1);

(iii) l(c2,g2); (iv) u(c2,g2); (v) l(c3,g3); (vi) u(c3,g3); (vii) l(c4,g4); (viii) u(c4,g4); (ix) l(c5,g5); (x) u(c5,g5); (xi) l(c6,g6); (xii) u(c6,g6). This convex hull region indicates that for any given specific value of GPA or CGPA [or an equivalent sGPCCU value] there exists an interval domain representing possible values for the corresponding SAPI values [or an equivalent sRSCCU values]; and that for any given specific SAPI value [or an equivalent sRSCCU value] there exists an interval domain representing possible values for the corresponding GPA or CGPA values [or an equivalent sGPCCU values].

This generic system characteristic has been already well illustrated in the two cases of typical example scenarios presented earlier and represented in Figure-1.1 and Figure-1.2 for the case of a 10-point grade points scale, and again in Figure-1.3 and Figure-1.4 for the case of a 5-point grade points scale. Similar plots of the convex hull region can be constructed for any given specific mapping scheme used for the conversion of the raw scores to letter grades.

The existence of a convex hull which is a region of finite area allows for multitudes of possible scenarios wherein the three kinds of chaotic system-intrinsic phenomena (as reported in the earlier part) can be observed, all leading to chaotic system biased information loss or corruption.

An Interlude - with the Author

Great! Let us look back for a moment before marching ahead. We have gone through a very meticulous analysis of the letter grading system, in order to understand how the intrinsic features in a system design can influence the observed overall system behavior.

The most critical design feature in the case of the letter grading system is the use of a *very crude grade point scale*, with an *extremely poor scale precision*, which introduces chaotic information loss/corruption, as has already been observed earlier.

The *aggregation process* itself, used in the computation of GPA and/or CGPA as the weighted average of the individual course grade points, is yet another source of chaotic information loss/corruption.

Also, the use of the Course Credits/Units (CCU), as *multiplicative weighting factors*, in the computation of the GPA and/or CGPA has been shown to be another source of chaotic information loss/corruption.

It is a mathematical fallacy *to claim that the result of any computation can have a better precision level than that of the input data set.* This is a serious flaw that is intrinsically implied in the LGS system design philosophy itself, since (i) *a significantly higher precision level is mysteriously presumed to have been achieved in the final numbers of GPA or CGPA, etc.;* while also (ii) wrongly *presuming that the teacher's precision in evaluation is rather poor, limited to classification into possibly only a handful of distinct categories*. These two factors together constitute the root cause for all the undesirable chaotic system characteristics of the LGS system.

In summary, the highly complex counter-intuitive and undesirable chaotic behavior of the LGS system is only due its poor system design. Therefore, a complete re-design of the system is required in order to overcome the 'APES-USA' problem. A proposed alternative system design is presented in the next part.

PART - III

A Proposed Solution

SAPES

Students
Academic
Performance
Evaluation
System

3.1 Introduction to the SAPES Environment

Now, let us apply the ideas and concepts developed earlier, to the context of the Students Academic Performance Evaluation System (SAPES) that we like to design. When a teacher designs a syllabus for a course, there is an expectation in the mind of the teacher, regarding the extent or degree of academic proficiency that a student is expected to achieve, by studying through that specific course. Whether such an expectation is made explicit or otherwise remains implicit, is immaterial at this point for our analysis and design. The extent or degree of academic proficiency that a student achieves by/through the learning process, is evaluated or measured by the teacher as an evaluator. In order to make these evaluations as objective as is realistically possible, the academic proficiency achieved is in fact measured through the evaluation or measurement of the academic performance of a student in a (seemingly continuous) sequence of tests or examinations or other evaluation schemes. The performance level of each of the students, in each of such tests is evaluated by the teacher, and the combined result of such evaluations, is considered to be an acceptable measure of the academic proficiency achieved by the student in that course, during a particular academic term.

Even without considering the situation for comparisons among the various possible student performance levels, in any

subject, it is possible for a teacher to imagine having an absolute measurement scale, with two fixed points, namely the *NULL Performance Point*, and the *FULL Performance Point*; usually corresponding to the 0% point and the 100% point on a percentage scale. Both these fixed points are usually very well identifiable, for a teacher, and is directly defined in terms of the extent of academic proficiency that a student is expected to achieve by the study of that course material. The tests/exams/etc can be designed carefully, to be able to measure, as well as to distinguish the various possible performance levels of the students.

One need not unnecessarily bring in the complications of possible non-linearities in the scale for this purpose. The non-linearities occur (if at all) mainly because of poorly specified syllabus or poorly designed tests; and should not be considered as intrinsic in the *design* of the *measurement system structure and mechanism*. It is clear that any well designed test (or a sequence of tests, or some other evaluation scheme) must incorporate an appropriately selected set of questions such that through such an evaluation, the true level of academic proficiency of every student in the class could be measured in an *unbiased, fair and reliable measurement scheme*. For this purpose, it is necessary that the test questions must evenly span the breadth and depth of the entire syllabus of the particular course. Also, it is certainly desirable in that process of evaluation, to be able to clearly distinguish the various possible performance levels, without much ambiguity. It is certain that a poorly designed test may end up being a curse to an academically proficient student, one

who may not get an opportunity to distinguish oneself through one's performance (in such a test) from every other student; whereas it can even be a boon to another one who is in fact not so proficient, but rather happens to be just simply lucky in this particular instance.

No system can be made completely fool-proof in this respect, except by the incorporation of appropriate checks and balances, in the system policies and procedures, or the rules and regulations; although we will not get into any further details on all these and many related aspects of the outer environment of the system rather than the system design per se.

3.2 Student Academic Performance (SAP) Scale

It is proposed that the measurement scale must be given the central focus in the design of the Students Academic Performance Evaluation System SAPES. *The Scale Quantum Unit Interval Domain Size* SQUIDS *is fixed to be equal to* UNITY, and is used as the most fundamental basis, as a single, universally common, unique standard unit of measurement, in the design of a unique SAPES measurement scale, which we can refer to as the *SAPES Scale*. The Scale Range can be represented by identifying the Scale Minimum Point SminiP and the Scale Maximum Point SmaxiP, which usually (but not necessarily, in general, as we shall see later) may correspond to the 'NULL' Point and the 'FULL' Point respectively.

The *Student Academic Performance* (SAP) *Scale Range* must correspond to the entire spectrum of possible academic performance levels, between the Scale Minimum Point SminiP and the Scale Maximum Point SmaxiP. Usually, the two well defined *Scale Fixed Points*, the *NULL Performance Point* NPP and the *FULL Performance Point* FPP define the effective SAP Scale Range.

The SAP Scale can be so designed for each of the different courses, that the Null Performance Point is common for all these different course specific SAP Scales. This has an intuitive appeal as well, since it takes the same (zero or nil) extent of achievement of academic proficiency in order to perform at the NULL-Performance level, irrespective of whichever be the specific course being considered. This universally common *Null Performance Point* NPP is the common zero point for every SAP Scale, irrespective of the course, and is therefore called the *SAP Absolute Zero Point*, in the *SAPES Measurement Scale System*.

With this unique common SAP absolute zero point, and with the SQUIDS as the common fundamental unit of SAP scale measurements; it is clear that the only variations possible in the different SAP scales corresponding to the different courses, is the location of the course-specific FULL Performance Point cFPP, and therefore the course-specific SAP Scale Range for each of the particular courses. With this design, it is possible to have a *single, universally common, unique Standard SAP Scale*, which we will refer to as the *SAPES Scale*. That is, the SAP scales corresponding to the

various different courses are in fact identical to one another, except for the location of the FULL Performance Point cFPP, or in other words, except the SAP Scale Range, corresponding to each of these courses, in the *SAPES Scale*.

The location of the course-specific cFPP corresponding to a specific course, on the SAPES Scale [or equivalently, the number of Scale Quantum Unit Interval Domains between NPP and FPP of the SAP scale corresponding to that specific course, which is also the SAP Scale Range for that course] is the course-specific maximum value for the raw-score SAPs that a student can earn in that course; defined as the course-specific *SAP Credits* cSAPc for that course, expressed in the "SQUIDS" unit of measurement. Therefore, for any course, 0 =< SAPs =< cSAPc.

3.3 Mapping between existing and proposed systems

For the sake of easy understanding, a mapping is presented here, between (the grade points scale of) the existing LGS system in the APES environment and the proposed SAP Scale in the SAPES environment; corresponding to any specific course. The course specific SAP Credits cSAPc for a course can be considered to be somewhat related to the currently prevalent concept of the Course Credits/Units CCU in the existing letter grading system. Although the two are *numerically not the same*, one can in fact draw a simple

relationship between the two, by trying to match the expected/desired behavior of the existing APES system, to what can in fact be actually observed with the new system design. With this in mind, one can consider the course-specific SAP Credits cSAPc to be simply *proportional* to the number that refers to the Course Credits/Units CCU. For this purpose, we define the *Credit/Unit SAP Score Points* CUSPs to be the course-specific SAP Credits cSAPc that is assigned to a one-credit course with CCU of unity; that is, cSAPc = CUSPs * CCU. A typical value may be a CUSPs value of 100 SQUIDS per credit [i.e., CUSPs is a proportionality constant.] Then, a 4-credit course with CCU=4, will be considered to carry a cSAPc value of 400 SQUIDS. Therefore, the cFPP for a 4-credit course would be a point on the SAPES Scale indicated by its cSAPc value of 400 SQUIDS. Similarly, a 3-credit course with CCU=3, is considered to carry a cSAPc value of 300 SQUIDS. Therefore, the cFPP for a 3-credit course would be a point on the SAPES Scale indicated by its cSAPc value of 300 SQUIDS. In this example, the CUSPs value of 100 SQUIDS per course credit unit gives rise to a precision level of the SAPES Scale [that is one SQUIDS] to correspond to *one hundredth of one* CCU. We may as well consider an alternative scenario, where one CCU is assigned 1000 CUSPs instead of 100 CUSPs, in order to achieve a tenfold improvement in the scale precision.

It may be useful to note here that the above mentioned SAPES Absolute Zero Point SAZP remains as the same common fixed point for even a SAP Scale that may need to be designed to

incorporate even negative values for possible SAP raw score measurements. For example, in the case of the written aptitude test scenario mentioned earlier, one may design a special SAP scale with the Scale Minimum Point SminiP as a point on the SAPES Scale that is marked as –100 SQUIDS (i.e., -25 Marks), corresponding to the lowest possible performance level; and the Scale Maximum Point SmaxiP as a point on the SAPES Scale that is marked as +400 SQUIDS (i.e., +100 Marks), corresponding to the highest possible performance level; with the SAPES Absolute Zero Point SAZP as the universally common NULL Performance Point. The SQUIDS value for such a SAP Scale would then correspond to 0.25 Marks as usually awarded. That is, each correct answer gets a positive score of four SQUIDS (+1 Mark), and each wrong answer gets a negative score of one SQUIDS (–1/4 Mark) in the SAPES Measurement Scale.

So, the value of any score as usually awarded in Marks can be easily and directly converted to the corresponding SAPs value expressed in SQUIDS units on the SAPES Measurement Scale, by just multiplying by a factor of four, since one Mark corresponds to four SQUIDS (for example, refer to conversion of Absolute Temperature figures expressed in degrees Rhankine and degrees Kelvin by the appropriate multiplication factor; unlike in the case of conversion between degrees Centigrade and degrees Fahrenheit). Thus, the SAPES Measurement Scale provides a universally common structural basis as a rational and systematic framework to combine/compare the scores obtained in (pertaining to) various distinctly different environments (like, for example, the scores

obtained in Written Aptitude Test, along with that contained in the Transcripts or Grade Reports, etc).

Every teacher is expected to award a valid SAPs *Raw Score* simply called as the SAP-score SAPs, for each of the students, in each of the courses that a student "registers for cSAPc credit", which has to be a positive integer in the range from zero to a maximum possible value given by an integer number equal to the cSAPc for that specific course. The individual teacher may be given the freedom as to how to arrive at a valid SAP-score SAPs, which is an appropriate, unbiased, fair, and reliable measure of the student academic performance; possibly using a sequence of tests, etc; about which, some general guidelines may be offered by the department or school or college or university. No intermediate conversions or transformations are to be performed, which would affect the information content represented by these SAPs raw scores. These SAPs scores associated with the various courses (corresponding to a semester, or an academic term/session, or the entire degree program) are *just simply added* together to compute a sum $\sum$SAPs, that is the total of all the SAPs scores, or even possibly a running cumulative if needed. Finally the *Student Academic Performance Index* (SAPI) is simply the ratio, ($\sum$SAPs)/($\sum$cSAPc). Here, $\sum$cSAPc is the sum total of the cSAPc values of all the courses that are to be accounted for, as per the graduation requirements, and other applicable policies and procedures or rules and Regulations of the school. The Student Academic Performance Index, SAPI may either be reported as such, explicitly giving the two numbers $\sum$SAPs and $\sum$cSAPc; or as a

fraction or even as a percentage value representing that ratio. Similarly, the *cumulative Student Academic Performance* cSAPI of a student for the entire program of study, can also be computed. Some of these various figures may as well be reported as a percentage; maybe for convenience in communication, and possibly to check for the individual course requirements, like minimum for Pass, course Distinction, etc, if at all such things do matter. However, all the original evaluations or measurements, all the intermediate representations, and also all the computations, are to be performed conforming to the SAPES Scale, as explained, so as not to introduce any undesirable system characteristics.

Now, let us get into further explanations on the above proposed system design. Having a single unique Standard SAPES Scale, as the central system component, and having fixed the fundamental/basic standard unit of measurement to be the SQUIDS unit, in the design of that measurement scale, means that every teacher uses a mark/score of UNITY (and not any fractional mark/score), as the smallest quantity of mark/score for the purpose of any evaluation. In special situations requiring a non-unit value for the SQUIDS, like in the case of the written aptitude test mentioned earlier in this paper, appropriate conversion has to be performed as illustrated there, in order to express the SAPs value in the *Universally Common Unique Standard SAPES Absolute Scale* with the Absolute Scale Precision (SQUIDS) of UNITY and the universally common SAPES Absolute Zero Point being defined as one of the Scale Fixed Points at NULL Performance Point NPP. This standardizes the

entire world of Student Academic Performance Evaluation System - SAPES environment - into a manageable entity for the purpose of design, analysis as well as implementation, without in any way causing undue restrictions or any inconveniences to the teachers, or the students or any other personnel who is required to work with the system. All unacceptable ambiguities in system definition, as well as all undesirable system complexities as exhibited in the LGS system behavior, are clearly avoided by this approach.

3.4 Some General Comments on SAPES Environment

Now, let us address some of the possible questions that could be raised by the proponents of a "distribution based" evaluation scheme. It may be generally felt that the raw scores and therefore the SAPI index of student academic performance evaluations system described above, can be to a large extent possibly dependent on the relative leniency or strictness of the specific teachers or evaluators, except in special situations where the tests themselves are standardized, with objective multiple choice questions, each associated with possibly one correct answer choice, or a best answer choice, to choose from the multiple choices provided, in which case the Scale Quantum Unit Interval Domain Size (SQUIDS) unit of measurement thereby gets defined clearly (although without a deliberate conscious design effort towards this specific aim of developing an unambiguous standard SQUIDS unit for

measurement). However, the solution to such a problem of possible variability (dependency on the teachers' relative leniency levels) is certainly not to go for the "distribution based" evaluation scheme, which would exhibit an almost equal extent of variability and an explicit dependency on the distribution of the raw scores in a class, which again is directly dependent on the individual teacher's leniency level, etc. Also, the fact that all the undesirable system characteristics like the CSBAAD, CSBSED, and CSBRRI still persists in this scheme. The only difference is that the location and extent/degree of the discontinuous system behavior, etc. would now be dependent on the quantization interval domain boundaries which get defined by each of the individual teachers for each course at each instance as needed, and hence can be quite ad hoc in nature. So, in addition to all the above mentioned (and analyzed) undesirable system characteristics, this distribution based evaluation scheme also suffers from an added dimension of undesirability, due to this ad hoc system design, and therefore certainly requires one to look for a different alternative system design, anyway.

The best solution to the problem of *teacher-to-teacher variability*, or rather the problem of *dependency on the teachers' leniency/strictness levels*, is in fact to provide appropriate information (in the grade-reports or now we may call it the 'SAPs Report') about such dependency. One can also possibly think of an extremely ideal situation, wherein the entire distribution of each of the SAPs scores is provided every time a specific SAPs score is reported. But, this solution can of course become quite enormous to be of practical

value. At least to start with, one can certainly think of incorporating in the overall *SAPES* system, a policy to include some information regarding the *Class Academic Performance Statistics* (CAPS); specifically the five statistical parameters, namely, CAPSsize - the *Class Size* (number of students in the class), CAPSmini - the Class Minimum SAPs score, CAPSmaxi - the Class Maximum SAPs score, CAPSmean - the Class Mean (Average) SAPs score, and possibly CAPSmedi – the Class Median SAPs score; associated with each of the courses for which a SAPs score is reported. Such statistical data would certainly provide the right perspective to an interpretation of the specific SAPs scores.

The reporting and availability of the data associated with the class performance statistics for each of the courses and also the overall school performance statistics can possibly be used for regulating the menace of grade inflation, although we do not intend to get into those details here, since it will lead to unnecessary diversion from our main course of discussion.

Note that any normalization or any transformation based on whatever standard distribution curve will only result in information loss/corruption, as mentioned earlier. How exactly these additional statistical data, associated with each of the SAPs scores are to be utilized in a systematic analysis of a set of student SAPs Reports is a subject for possibly a separate study. The only point to be stressed here is that, each of the individual SAPs scores must not be tampered with, by any kind of normalizations or transformations or whatever,

and must be left as such. Any further analyses on the data supplied, can always be carried out based on the actual need for such analyses.

A teacher who still decides/wishes to use fractional marks/scores in any evaluation or measurement, in this new SAPES environment, may certainly be allowed to do so, as long as the SAPs score that is reported in fact corresponds to the SAP Scale for that specific course, with the cSAPc value as the maximum possible value for any SAPs score thereof. The only comment here is that such a teacher may simply be over-exerting, with or without some well defined deliberate intent thereof. On the other hand, if the smallest chunk of mark/score that a teacher uses is significantly larger than unity (that is, one SQUIDS unit) then it may seem to indicate that this teacher may not probably be effectively utilizing the available scale precision for the purpose of an unbiased fair and reliable evaluation/measurement of the students academic performance levels, contrary to the expectations of both the students community as well as the employers, etc. Usually, this enables the department, or the school (or the college, or the university) to evolve appropriate policies & procedures, or rules & regulations, for establishing the necessary checks and balances in the overall SAPES environment. However, it is to be noted here that, there may be situations corresponding to some courses/subjects, wherein the very nature of the subject material and the possible methods available for measurements of a student's academic performance in such situations, impose certain constraints on the measurable precision levels thereof. It is important to note here that the available scale precision

(SQUIDS) must be at least as much as or even better than that required in any/every measurement scenario that uses such a measurement scale. In other words, while using any measurement scale, the best achievable measurement precision cannot be better than that provided by the measurement scale. However, if situation demands, we may always be able to use the same measurement scale for actual measurements requiring relatively less precision than what is provided by the measurement scale.

Also, it is interesting to note that the SAPES measurement scale, and the system design in principle, does not necessarily prevent a teacher from assigning a SAPs score the numerical value of which is larger than that of the corresponding cSAPc credits, and therefore may at the outset seem to be invalid as per the earlier discussion. However, even such exceptional situations, indicating that the academic performance level of that particular student in that specific course has been evaluated to be beyond the expected standard full performance level, may be allowed. Here, and in such similar situations, it is of course desirable to require certain appropriate specific approvals from the department and/or the school (and/or the college and/or the university) so as to provide appropriate checks and balances, in terms of policies & procedures, or rules & regulations, incorporated into the overall system environment. It is quite a simple matter to implement such a possible policy/regulation on the possible validity of the SAPs score, if so required, to be within some appropriately specified limits identified by the Scale Minimum Point SminiP and the Scale Maximum Point SmaxiP.

As a final note, it is clear that the same SAPES Environment may very well be used even if someone (teacher or department or school or whatever) decides to go with the idea of "relative evaluation" or "distribution based evaluation"; that is equivalent to what has been usually referred to as "grading on the curve" in the existing LGS system. The availability of an absolute scale of measurement can certainly be made use of even when a situation may in fact require relative evaluation, or even if any nonlinear transformation (or even any data dependent transformation) is required. However, it is usually well accepted that the use of any nonlinear scale for measurement may introduce unnecessary complexity or may even cause loss of information, and therefore to be considered as undesirable, especially when such nonlinearities cannot generally be justified. The concept of distribution based evaluation, can introduce significant data dependent nonlinearities, and may possibly be associated with the uncertainties in the parameters of such distribution, thus introducing another dimension of complexity, in addition to causing information loss or rather an irrecoverable information mutation or corruption, that is an intrinsic characteristic in the very design of such systems, as we have already analyzed earlier.

3.5 Summary of the SAPES system design

A common unique universal SAPES measurement scale is designed as an linear absolute scale; with the SAPES Absolute Zero Point SAZP fixed at the Null Performance Point NPP as one of the Scale Fixed Points marked as 0 SQUIDS; and the Scale Quantum Unit Interval Domain Size SQUIDS defined as Unity (one SQUIDS). The second Fixed Point required to define a complete measurement system to enable the measurement of performance levels in any specific course, is a course specific Full Performance Point cFPP. The course specific cFPP is determined by the course specific SAP Credits cSAPc assigned to the course. This course specific cSAPc is the usual maximum value of the raw-score SAPs that a student can earn in that specific course.

To facilitate a smooth and easy mapping between the existing LGS system and the proposed SAPES system, the concept of course specific credit unit score points CUSPs has been conveniently introduced, which works as a simple proportionality constant multiplier to convert the course specific credit units CCU to the course specific SAP Credits cSAPc; that is, cSAPc = CUSPs * CCU. For example, a nominal value for CUSPs is 100, although it is not unreasonable to assign a value like 1000 in order to enhance the precision of the measurement system based on the precision expected or required in the final reported figures for SAPI and/or cSAPI.

No teacher is required to use any fractional numbers for raw scores, the minimum quantum for raw score being 1 SQUIDS. All aggregate values are computed by simple arithmetic addition, all averages values are

computed as simple arithmetic mean values. For example, a cumulative or overall Student Academic Performance Index cSAPI can be conveniently defined as the ratio : cSAPI = $(\sum \text{SAPs})/(\sum \text{cSAPc})$.

It is recommended that SAP Report (in place of the Grade Report) must also include some appropriately selected *Class Academic Performance Statistics* (CAPS) [and possibly also the school level performance statistics data]; especially the five essential statistical parameters: that is, (i) CAPSsize - the *Class Size* (number of students in the class), (ii) CAPSmini - the Class Minimum SAPs score, (iii) CAPSmaxi - the Class Maximum SAPs score, (iv) CAPSmean - the Class Mean (Average) SAPs score, and possibly (v) CAPSmedi – the Class Median SAPs score; associated with each of the courses for which a SAPs score is reported. Such statistical data would certainly provide the right perspective to an interpretation of the specific SAPs scores. One can also possibly think of an extremely ideal situation, wherein every time a specific SAPs score is reported, the entire distribution of the SAPs scores for that course is also provided. This may however require the appropriate support of the modern MIS & data base technology.

* * *

An Afterword

The APES problem has been well studied. The letter grading system certainly needs to be replaced by a better alternative. The proposed alternative system design 'SAPES' can in fact be considered for potential deployment / implementation.

However, as is always the case, any 'change' has to face the inertial resistance, usually encountered in real life. All reforms encounter resistance at every step. Reforms and even revolutions come about, just as a chain reaction in an explosion, only because of the preceding potential build up - as if waiting for a trigger so as to explode.

Academia in general, and student issues in particular, are the least prone to any potential build up, whether for good or otherwise - docility being the prime reason, priority concerns and cycle time being the other possible reasons. Therefore, the 'change' that is sought after must necessarily be of an evolutionary nature rather than a revolutionary one. This requires a persistent and consistent application of well focused concerted effort by everyone concerned, in the intellectual process of convincing and/or getting convinced!

If you are not yet convinced, begin by being open to allow a possible 'change' to happen from within, expose yourself to some of these novel ideas & concepts. When once you are convinced, maybe you can initiate a quiet and peaceful revolution through a collegial process of trying to convince your close associates.

ABOUT THE AUTHOR

Since 1991, Dr. Keshava PRASAD Halemane has been working as Professor of Systems Analysis and Computer Applications in the Department of Mathematical And Computational Sciences at National Institute of Technology Karnataka Surathkal, which is also his Alma Mater, earlier known as Karnataka Regional Engineering College Surathkal.

Prasad has an impressive academic credentials. He secured the university first rank as well as the state level overall highest percentage of aggregate marks, considering all the examinations of all the semesters of the undergraduate engineering B.E. degree program, across all the universities of the Karnataka State in 1972, and was awarded the 'Karnataka State Gold Medal for The Best Engineering Student of The State'. His *doctoral research work* in CMU in the area of '*optimum design of engineering systems that accounts for the uncertainties in parameter values*' has effectively revitalized the research investigations in that area, by his novel initiation in *opening up an entirely new approach* which has now been *well recognized and followed,* as the *most rational and systematic conceptual advancement* in that field.

In addition to his long tenure in academia, Prasad has extensive experience in various fields - including his initial exposure to the manufacturing industry, corporate as well as national R & D work, both in USA as well as in India. He has shown convincingly, that he is very quick and efficient in learning new things and attaining expertise in new areas; and has felt very comfortable in dealing with the variety of diverse work situations that he has gone through, being distinctly outstanding among his colleagues and co-workers in every organization that he worked for. Prasad has demonstrated excellent leadership in every situation, and has been singularly conspicuous among his colleagues, indeed as a visionary genius.

Above all these, Prasad is endowed with a rare combination of an extremely sharp intellect (depth-in-thinking) along with a highly sensible mind (balance-in-judgement) and a humane heart (breadth-in-feelings). Also, he is noted for his integrity of character and pleasant disposition. He has a special talent for *teaching* in its broadest sense, and is an exemplary (inspiring) teacher indeed, although (or maybe, because) he prefers to consider himself as a student all through life: "Life, the best teacher; and Living, the best Learning"!

To interact/participate/contribute visit: http://apes-usa.blogspot.in/